

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 16:07:29 ; Search time 13988 Seconds

(without alignments)  
11511.316 Million cell updates/sec

Title: US-09-720-383C-9

Perfect score: 3936

Sequence: 1 cttctccctcgtcgtcgcgg.....ctgagtcgttttctgtaaaa 3936

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3935	100.0	3968	8	AF200531	AF200531 Zea mays
2	2543	64.6	3538	8	AF200530	AF200530 Zea mays
3	2241	56.9	3812	8	AF200532	AF200532 Zea mays
4	1531.4	38.9	3847	8	AF304374	AF304374 Nicotiana
5	1509	38.3	3255	6	AX507835	AX507835 Sequence
6	1489.8	37.9	3444	8	AF062485	AF062485 Arabidops
7	1434.2	36.4	3771	8	AF027173	AF027173 Arabidops
8	1431.8	36.4	3255	6	AX505864	AX505864 Sequence
9	1429.4	36.3	3828	6	AR267558	AR267558 Sequence
10	1429.4	36.3	3828	6	AX030944	AX030944 Sequence
11	1429.4	35.0	3725	8	BD022677	BD022677 Manipulat
12	1378.4	35.0	3725	8	AF200526	AF200526 Zea mays
13	1364	34.7	3752	8	AF200525	AF200525 Zea mays
14	1335.4	33.9	171376	8	AC104487	AC104487 Oryza sat
15	1332.4	33.9	3763	8	BT008654	BT008654 Arabidops
16	1330.8	33.8	3673	6	AR267560	AR267560 Sequence
17	1330.8	33.8	3673	6	AX030948	AX030948 Sequence
18	1330.8	33.8	3673	6	BD022679	BD022679 Manipulat
19	1326.6	33.7	3603	6	AR267557	AR267557 Sequence
20	1326.6	33.7	3603	6	AX030942	AX030942 Sequence
21	1326.6	33.7	3603	6	BD022676	BD022676 Manipulat
22	1321.4	33.6	3676	8	AF200529	AF200529 Zea mays
23	1301.8	33.1	3723	8	AF150630	AF150630 Gossypium
24	1298.2	33.0	3222	8	AX652952	AX652952 Sequence
25	1294.8	32.9	150778	2	AP004298	AP004298 Oryza sat
26	1289.4	32.8	3745	8	AF200528	AF200528 Zea mays
27	1279.4	32.5	3795	8	AF200533	AF200533 Zea mays
28	1279.4	32.5	3799	6	AX338680	AX338680 Sequence
29	1272.4	32.3	3229	8	BT002335	BT002335 Arabidops
30	1269.2	32.2	3614	6	AR267559	AR267559 Sequence
31	1269.2	32.2	3614	6	AX030946	AX030946 Sequence
32	1269.2	32.2	3614	6	BD022678	BD022678 Manipulat
33	1269.2	32.2	3682	8	AF027174	AF027174 Arabidops
34	1268.4	32.2	3264	6	AX653232	AX653232 Sequence
35	1250.6	31.8	3395	8	AF081534	AF081534 Populus x
36	1198	30.4	3081	8	AF088917	AF088917 Arabidops
37	1196.8	30.4	3355	8	AY139754	AY139754 Arabidops
38	1196.4	30.4	3081	8	BT004543	BT004543 Arabidops
39	1192.6	30.3	3277	8	AY095297	AY095297 Populus t
40	1192.4	30.3	3181	8	BT006111	BT006111 Arabidops
41	1192.4	30.3	3377	8	BT005710	BT005710 Arabidops
42	1184	30.1	3168	8	AF458083	AF458083 Arabidops
43	1171.2	29.8	2828	8	AF200527	AF200527 Zea mays
44	1170	29.7	2830	6	AX338676	AX338676 Sequence
45	1163.4	29.6	130300	8	AP005824	AP005824 Oryza sat

#### ALIGNMENTS

RESULT 1  
AF200531 3968 bp mRNA linear PLN 31-AUG-2000  
LOCUS  
DEFINITION Zea mays cellulose synthase-7 (Cesa-7) mRNA, complete cds.  
ACCESSION AF200531  
VERSION AF200531.1 GI:9622885  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
1 (bases 1 to 3968)  
Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,  
AUTHORS Xocconostle-Cazares,B. and Delmer,D.F.

Pred. No. is the number of results predicted by chance to have a

TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene family  
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)  
MEDLINE 20398328  
PUBMED 10938350  
REFERENCE 2 (bases 1 to 3968)  
AUTHORS Dhugga, K.S. and Helentjaris, T.G.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA  
FEATURES  
source  
1. .3968  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"  
1. .3968  
/gene="Cesa-7"  
143. .3403  
/gene="Cesa-7"  
/note="Cesa-7"  
/codon\_start=1  
/product="cellulose synthase-7"  
/protein\_id="AAF89967.1"  
/db\_xref="GI:9622886"  
/translation="MEASAGLVAGSHNRNELVIRRDGDPKPPREONGOVCOICGD  
DVGLAPGDPFVACNECAPVYCRDYEYERREGTQNPQCKTRYRLKGCQRYTGDEB  
EDGVDLDNEFNWDGSDSVAESMLYGHMSYGRGDPNGAFOALPNVPLTNGO  
MVDDIPPEQHALVPSFMGGGKRIHPLPYADSLPVQPRSMDSKDLAAYGSSVAWK  
ERMENWKQROERHQGTGNDGGDDDDADLPLMEARQQLSRKIPLSSQINPYRMII  
IRLVVLGFFFHYRVMPVNDAFALMLISVCEIWFAMSWILDQPKWFIERYTYLD  
RLSLRFDKEGQSQLAPIDFVSTVDPLKEPLVTTNTVLSILSDVPVDKVSCTVSD  
DGAAMLTFEALSTSEPAKWPVPCRYNIEPRAPEWYFOQKIDYLDKDKVAANFVRER  
RAMKREYEEFKVRINALVAKAQKVPBEGWMTQDGTWPENNVDRHPGMIQVFLQSGSG  
LDCEGNELPRLVYVSREKRPYNNHKKAGAMNALVRSAYLTNAPYLNLDCHYINN  
SKAIKEAMCFEMDPLLGKVCYVQFQRPFDGIDRHDRYANRVFEDINMKGLDIG  
PIYVGTGCVFRQALVGYDAPKTKPPSRTCNCWPKWCFCCCFGNRKQKTKPKTE  
KKKLLFPKKEENQSPAYALGEIDEAAPAENKAGIVNQKLEKFGQSSVFTYSTLL  
ENGTLKSASPASLKEALHVISCGYEDKTWKEIGWISYVTEIDLTFKMHGHW  
RSYICIPKRVAFKGSAPLNSDLHQLRWALGSIEIFFSNHCPLMYGGGLFLER  
FSYINSIVPMWTSIPLAICTLPAICLLTGKFTPELNNVASLWMSLFCIFATSL  
EMRWSGVGIDDMWRNEQFWVIGVSSHLFAVFOGLVIAVDTSFTVTSKGDDEEF  
SELYTFKWTLLIPPTILLNFIQVAVGSNAINNGYESWGPLFGKLFPAFVIVHL  
YPLKGLVGRNRTPTIYVWSILASIFSLWVRIDPLAKDGPLLECGLDGN"  
BASE COUNT 1022 a 837 c 1018 g 1091 t  
ORIGIN  
Query Match 100.0%; Score 3935; DB 8; Length 3968;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3935; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 TTCTCCCTCGTCGGTGGCGCGGCTCGGCGTTCGGTGAAGAAACCACTCGGGGA 61  
DB 1 TTCTCCCTCGTCGGTGGCGCGGCTCGGCGTTCGGTGAAGAAACCACTCGGGGA 60  
QY 62 TGAAGATCTGCTGTAGATGAGAGACTACGGTCAGTATCCTTGCCTTGTGCGCG 121  
DB 61 TGAAGATCTGCTGTAGATGAGAGACTACGGTCAGTATCCTTGCCTTGTGCGCG 120  
QY 122 CGAAGTGAAGGAGGAGGAGCGATGAGCGCGCGGCTGTGCGCGGCTCCACAA 181  
DB 121 CGAAGTGAAGGAGGAGGAGCGATGAGCGCGCGGCTGTGCGCGGCTCCACAA 180  
QY 182 CCGCAACGAGCTCGTCATCCGCGCGACCGCGATCCCGGGCGAAGCCGCGCGGA 241  
DB 181 CCGCAACGAGCTCGTCATCCGCGCGACCGCGATCCCGGGCGAAGCCGCGCGGA 240  
QY 242 GCAGAACGGGCAAGGTGTGCGAGATTGCGGCGACGACGTCGGCCTTGCCCCGGCGGGA 301  
DB 241 GCAGAACGGGCAAGGTGTGCGAGATTGCGGCGACGACGTCGGCCTTGCCCCGGCGGGA 300

QY 302 CCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCGGCTCTGCGGGGACTGTACGAATACGA 361  
DB 301 CCCCTTCGTGGCGTGCAACGAGTGCGCCTTCCCGGCTCTGCGGGGACTGTACGAATACGA 360  
QY 362 GCGCCGGGAGGCGACGACGAACTGCCCCCAGTGCAGAACTCGATACAAGCGCCTCAAGGG 421  
DB 361 GCGCCGGGAGGCGACGACGAACTGCCCCCAGTGCAGAACTCGATACAAGCGCCTCAAGGG 420  
QY 422 CTGCCAAGTGTGACCGGTGACGAGAGAGAGACGCGCGTGCATGACCTGGACAACGAGTT 481  
DB 421 CTGCCAAGTGTGACCGGTGACGAGAGAGAGACGCGCGTGCATGACCTGGACAACGAGTT 480  
QY 482 CAAGTGGAGCGGCCATGACTGCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAG 541  
DB 481 CAAGTGGAGCGGCCATGACTGCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAG 540  
QY 542 CTACGGCGGTGAGGTGACCTTAATGGCGCGCCACAAGCTTTTCCAGCTCAACCCCAATGT 601  
DB 541 CTACGGCGGTGAGGTGACCTTAATGGCGCGCCACAAGCTTTTCCAGCTCAACCCCAATGT 600  
QY 602 TCCACTCTCACCACGCGCAATGTGATGACATCCACGAGACGACGCGCTGT 661  
DB 601 TCCACTCTCACCACGCGCAATGTGATGACATCCACGAGACGACGCGCTGT 660  
QY 662 GCCTTCTTCATGCGGTGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCAG 721  
DB 661 GCCTTCTTCATGCGGTGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCAG 720  
QY 722 CTACCTGTGCAACCCAGGTCTATGGACCCATCCAGAGATCTTGCTGCATATGGGTATGG 781  
DB 721 CTACCTGTGCAACCCAGGTCTATGGACCCATCCAGAGATCTTGCTGCATATGGGTATGG 780  
QY 782 TAGTGTGCTTGGAGGACGAGATGGAATTTGGAAGCAGAGACAAGAGAGATGACCA 841  
DB 781 TAGTGTGCTTGGAGGACGAGATGGAATTTGGAAGCAGAGACAAGAGAGATGACCA 840  
QY 842 GACGGGAATGATGTTGTTGTTGATGATGTTGACGATGCTGATCTACCACTAATGATGA 901  
DB 841 GACGGGAATGATGTTGTTGTTGATGATGTTGACGATGCTGATCTACCACTAATGATGA 900  
QY 902 AGCAAGACAACAACCTGTCCAGGAAATTTCCACTTCCATCAAGCCAGATTAATCCATATAG 961  
DB 901 AGCAAGACAACAACCTGTCCAGGAAATTTCCACTTCCATCAAGCCAGATTAATCCATATAG 960  
QY 962 GATGATATCATATTTCGGCTTGTGTTTGGGGTCTTCCACTACCGAGTGATGCA 1021  
DB 961 GATGATATCATATTTCGGCTTGTGTTTGGGGTCTTCCACTACCGAGTGATGCA 1020  
QY 1022 TCCGGTGAATGATGATTTGCTTTGTGCTCATATCTGTTATCTGTGAATCTGTTTGC 1081  
DB 1021 TCCGGTGAATGATGATTTGCTTTGTGCTCATATCTGTTATCTGTGAATCTGTTTGC 1080  
QY 1082 CATGCTTGGATTCTTGATCAATTTCCCAAAGTGTTCCTTAATGAGAGAGACTTACCT 1141  
DB 1081 CATGCTTGGATTCTTGATCAATTTCCCAAAGTGTTCCTTAATGAGAGAGACTTACCT 1140  
QY 1142 AGACCGGCTGTCACTGAGGTTGCAAGGAGGCCAGCCATCTCAACTTGCTCCAATTGA 1201  
DB 1141 AGACCGGCTGTCACTGAGGTTGCAAGGAGGCCAGCCATCTCAACTTGCTCCAATTGA 1200  
QY 1202 TTTCTTTGTCAAGTACGTTGATCCCTTAAAGGAACCTCCTTGGTCAACAATACTGT 1261  
DB 1201 TTTCTTTGTCAAGTACGTTGATCCCTTAAAGGAACCTCCTTGGTCAACAATACTGT 1260  
QY 1262 TCTATCTATCCTTTGGGTGATTAATCCTGTGATAAGGTTTCTTCTATGTTCTGATGA 1321  
DB 1261 TCTATCTATCCTTTGGGTGATTAATCCTGTGATAAGGTTTCTTCTATGTTCTGATGA 1320  
QY 1322 TGGTGTGCAATGCTAACGTTTGAAGCAATTATCTGAACACATCTGAATTTGCAAGAAATG 1381  
DB 1321 TGGTGTGCAATGCTAACGTTTGAAGCAATTATCTGAACACATCTGAATTTGCAAGAAATG 1380  
QY 1382 GGTTCCTTCTGCAAAACGTAACAATATTGAACCTGCGGCTCCAGAGTGTAATCTTCAACA 1441

Db 1381 ||||| GGTTCTTTCTGCAAAAGGTATTTGAACCTCGCGCTCCAGAGTGGTACTTCCAACA 1440

QY 1442 GAAAGATAGACTACTTGAAGAACAAGGTGGCAAAACTTTGTTAGGGAGAGAGCAAT 1501

Db 1441 GAAAGATAGACTACTTGAAGAACAAGGTGGCAAAACTTTGTTAGGGAGAGAGCAAT 1500

QY 1502 GAAGAGAGAGTATGAGGAATTCAGGTGAGAAATCAATGCTTATGTTGCCAAAGCCAGAA 1561

Db 1501 GAAGAGAGAGTATGAGGAATTCAGGTGAGAAATCAATGCTTATGTTGCCAAAGCCAGAA 1560

QY 1562 AGTTCCTGAAGAAGGATGGAACAATGCAAGATGGAACCCCTGGCTGGAAACAATGTTGC 1621

Db 1561 AGTTCCTGAAGAAGGATGGAACAATGCAAGATGGAACCCCTGGCTGGAAACAATGTTGC 1620

QY 1622 TGATCATCTGGAATGATTCAAGTCTTCTGGCCAAAGCGAGGCTTGACTGTGAGGG 1681

Db 1621 TGATCATCTGGAATGATTCAAGTCTTCTGGCCAAAGCGAGGCTTGACTGTGAGGG 1680

QY 1682 AAATGAACCTGCCAGATTGGTTTATGTTCTAGAGAGAAACGACGAGCTATAACCATCA 1741

Db 1681 AAATGAACCTGCCAGATTGGTTTATGTTCTAGAGAGAAACGACGAGCTATAACCATCA 1740

QY 1742 TAAGAAAGCTGTGCTATGATGCATTGCTCCGAGTCTGCTGTACTAAACAATGCTCC 1801

Db 1741 TAAGAAAGCTGTGCTATGATGCATTGCTCCGAGTCTGCTGTACTAAACAATGCTCC 1800

QY 1802 ATATTGTTAACTTGGAATGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGC 1861

Db 1801 ATATTGTTAACTTGGAATGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAGC 1860

QY 1862 AATGTGTTTATGATGAGACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTCCCTCA 1921

Db 1861 AATGTGTTTATGATGAGACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTCCCTCA 1920

QY 1922 AAGATTGATGGGATTTGATCGCCATGACCGATATGCTAACCGGAATGTGCTTTTGA 1981

Db 1921 AAGATTGATGGGATTTGATCGCCATGACCGATATGCTAACCGGAATGTGCTTTTGA 1980

QY 1982 TATCAACATGAAAGGTTGGATGGTATTCAGGGTCCAATTTATGTTGTAAGTGTGT 2041

Db 1981 TATCAACATGAAAGGTTGGATGGTATTCAGGGTCCAATTTATGTTGTAAGTGTGT 2040

QY 2042 ATTTAGAAAGGCAAGCATTATATGTTATGATGCCCCCAAAACAAGACCACATCAAG 2101

Db 2041 ATTTAGAAAGGCAAGCATTATATGTTATGATGCCCCCAAAACAAGACCACATCAAG 2100

QY 2102 GACTTGCAACTGTGCGCCCAAGTGTGCTTTGCTGTGCTGCTTGGCAATAGGAAGCA 2161

Db 2101 GACTTGCAACTGTGCGCCCAAGTGTGCTTTGCTGTGCTGCTTGGCAATAGGAAGCA 2160

QY 2162 AAAGAAGACTACCAAAACCAAAACAAGAAGAAAAGTTATTTTCAAGAAAGAGA 2221

Db 2161 AAAGAAGACTACCAAAACCAAAACAAGAAGAAAAGTTATTTTCAAGAAAGAGA 2220

QY 2222 GAACCAATCCCTGCAATATGCTTTGGTGAATTTGACGAAGCTGCTCCAGAGCTGAGAA 2281

Db 2221 GAACCAATCCCTGCAATATGCTTTGGTGAATTTGACGAAGCTGCTCCAGAGCTGAGAA 2280

QY 2282 TGAAGAGCCCGGTATTGTAATCACAAAAAATTAGAAAAAGAAATTGGCCAATCTTCTGT 2341

Db 2281 TGAAGAGCCCGGTATTGTAATCACAAAAAATTAGAAAAAGAAATTGGCCAATCTTCTGT 2340

QY 2342 TTTTGTATCATCCACACTTCTCGAAGTGTGGAACCTTGAAGAGTCAAGTCCCTGCTTC 2401

Db 2341 TTTTGTATCATCCACACTTCTCGAAGTGTGGAACCTTGAAGAGTCAAGTCCCTGCTTC 2400

QY 2402 TCTTTGAAAAGAGCTATACATGTCATTAGTGTGTTATGAAGACAAGACAGACTGGGG 2461

Db 2401 TCTTTGAAAAGAGCTATACATGTCATTAGTGTGTTATGAAGACAAGACAGACTGGGG 2460

QY 2462 AAAAGAGATTGGCTGATCTATGATCAGTTACAGAAGATATTCTAAGTGTTCAGAT 2521

Db 2461 AAAAGAGATTGGCTGATCTATGATCAGTTACAGAAGATATTCTAAGTGTTCAGAT 2520

Db 2461 AAAAGAGATTGGCTGATCTATGATCAGTTACAGAAGATATTCTAAGTGTTCAGAT 2520

QY 2522 GCATGTGATGTTGGCGGTCAATTTACTGATACCTTAAACGGGTTGATTCAAAAGTTC 2581

Db 2521 GCATGTGATGTTGGCGGTCAATTTACTGATACCTTAAACGGGTTGATTCAAAAGTTC 2580

QY 2582 TGCACTCTGAATCTTTCAGATGCTTCAACAGGTGCTTGGTGGCTTGGGTCTAT 2641

Db 2581 TGCACTCTGAATCTTTCAGATGCTTCAACAGGTGCTTGGTGGCTTGGGTCTAT 2640

QY 2642 TGAGATCTTCTCAGCAATCATGCCCCCTTTGGTATGGGTATGGTGGCGCTGAAT 2701

Db 2641 TGAGATCTTCTCAGCAATCATGCCCCCTTTGGTATGGGTATGGTGGCGCTGAAT 2700

QY 2702 TTTGGAAGATTTTCTTACATCACTCCATCGTGTATCCTTGGACATCTATTCCTCTT 2761

Db 2701 TTTGGAAGATTTTCTTACATCACTCCATCGTGTATCCTTGGACATCTATTCCTCTT 2760

QY 2762 GGCCTACTGTACATTTGCCCTGCCATCTGTTTATTTGACAGGAAATTTATCACTCCAGAGCT 2821

Db 2761 GGCCTACTGTACATTTGCCCTGCCATCTGTTTATTTGACAGGAAATTTATCACTCCAGAGCT 2820

QY 2822 GAATTAATGTTGCCAGCCTGTGGTCAATGTCATTTTATCTGCATTTTGTCTACGAGCAT 2881

Db 2821 GAATTAATGTTGCCAGCCTGTGGTCAATGTCATTTTATCTGCATTTTGTCTACGAGCAT 2880

QY 2882 CCTAGAAATGAGATGAGTGGTGTGGAATTTGATGACTGCTGGTGAAGAAATGACAGTCTG 2941

Db 2881 CCTAGAAATGAGATGAGTGGTGTGGAATTTGATGACTGCTGGTGAAGAAATGACAGTCTG 2940

QY 2942 GGTCAATTGAGGTGTGCTCACAACCTCTTGTGCTGTGTTCCAGGACTTCTCAAGGTCAT 3001

Db 2941 GGTCAATTGAGGTGTGCTCACAACCTCTTGTGCTGTGTTCCAGGACTTCTCAAGGTCAT 3000

QY 3002 AGCTGTGTTGATACAAAGCTTCAACCGTGACATCAAAAGGTGAGATGAGAGGTTCTC 3061

Db 3001 AGCTGTGTTGATACAAAGCTTCAACCGTGACATCAAAAGGTGAGATGAGAGGTTCTC 3060

QY 3062 AGAGCTATATACATTCAAAATGACTACCTTATGATACCTCCACACTTGCTTCTATT 3121

Db 3061 AGAGCTATATACATTCAAAATGACTACCTTATGATACCTCCACACTTGCTTCTATT 3120

QY 3122 GAATTCATTTGTTGTGTGCTGGCGTTTCAATGCGATCAATTAACGGATATGAGTCATG 3181

Db 3121 GAATTCATTTGTTGTGTGCTGGCGTTTCAATGCGATCAATTAACGGATATGAGTCATG 3180

QY 3182 GGGCCCCCTCTTTGGGAAGCTATTTTGCATTTTGGGTGATGTCATCTTATCCCTT 3241

Db 3181 GGGCCCCCTCTTTGGGAAGCTATTTTGCATTTTGGGTGATGTCATCTTATCCCTT 3240

QY 3242 TCTCAAAAGTTTGGTTGAAGCAAAACAAGACACCAACGATTGTCATGCTGTGCCAT 3301

Db 3241 TCTCAAAAGTTTGGTTGAAGCAAAACAAGACACCAACGATTGTCATGCTGTGCCAT 3300

QY 3302 TCTGCTGGCTTCAATCTTCTCGCTCCTTGGGTTCCGATTGAATCTTCTTCCGGAAGA 3361

Db 3301 TCTGCTGGCTTCAATCTTCTCGCTCCTTGGGTTCCGATTGAATCTTCTTCCGGAAGA 3360

QY 3362 TGATGTCGCGCTCTTGAAGAGTGTGCTTTGGATTGCAACTRAGATGTCAGTCATCAGC 3421

Db 3361 TGATGTCGCGCTCTTGAAGAGTGTGCTTTGGATTGCAACTRAGATGTCAGTCATCAGC 3420

QY 3422 TCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTGTTGCCATATTCAGTGTG 3481

Db 3421 TCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGGTGTGTTGCCATATTCAGTGTG 3480

QY 3482 TGTAGATTAAGACATGAATGTCCTCAAGTTCTTTTGTATCCATGTTGAACCTTACTTAAT 3541

Db 3481 TGTAGATTAAGACATGAATGTCCTCAAGTTCTTTTGTATCCATGTTGAACCTTACTTAAT 3540

QY 3542 ATCTGAGAGATATACGCGGGAATAATGAGGCTGCGGCAATCCTTGTGAGTTGGCGGT 3601

Db 3541 ATCTGAGAGATATACGCGGGAATAATGAGGCTGCGGCAATCCTTGTGAGTTGGCGGT 3600



OY 3602 GGAATACAGCATATGCAAGTGTGATGTCGACATCTTTATATCTTGTGCGAATAT 3661  
|||||  
DB 3601 GGAATACAGCATATGCAAGTGTGATGTCGACATCTTTATATCTTGTGCGAATAT 3660  
OY 3662 AGATGGGCTGAGCCCAACAGCAAGGATTTTGAATCTGCACTGCTCCCGTGACAACTT 3721  
|||||  
DB 3661 AGATGGGCTGAGCCCAACAGCAAGGATTTTGAATCTGCACTGCTCCCGTGACAACTT 3720  
OY 3722 GGTCTCAATTAAGGAGGAGCAAGATGATCTGCGAGTGAACAGAGCAACTGACATTA 3781  
|||||  
DB 3721 GGTCTCAATTAAGGAGGAGCAAGATGATCTGCGAGTGAACAGAGCAACTGACATTA 3780  
OY 3782 TTTATGTATGCTGTTCTATTGAGGGCTGTTGATTAATGTTCTGTTAATAAGAAAA 3841  
|||||  
DB 3781 TTTATGTATGCTGTTCTATTGAGGGCTGTTGATTAATGTTCTGTTAATAAGAAAA 3840  
OY 3842 ACAGATATTAAGCATTAATCTATAGTTAATTAAGTAAATGCGCTGTTTTTTGTT 3901  
|||||  
DB 3841 ACAGATATTAAGCATTAATCTATAGTTAATTAAGTAAATGCGCTGTTTTTTGTT 3900  
OY 3902 GTGTACTGTATCATCTGAGTTGTTGTTGTAATA 3936  
|||||  
DB 3901 GTGTACTGTATCATCTGAGTTGTTGTTGTAATA 3935

RESULT 2  
AF200530 3538 bp mRNA linear PLN 31-AUG-2000  
LOCUS Zea mays cellulose synthase-6 (Cesa-6) mRNA, complete cds.  
DEFINITION  
AF200530  
VERSION AF200530.1 GI:9622883  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 3538)  
AUTHORS Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S.,  
Xoconostle-Cazares, B. and Delmer, D.P.  
TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene  
family  
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)  
MEDLINE 20398328  
PUBMED 10938350  
REFERENCE 2 (bases 1 to 3538)  
AUTHORS Dhugga, K.S. and Helentjaris, T.G.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred  
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA  
FEATURES  
Source Location/Qualifiers  
1. 3538  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/note="Identified and isolated from the maize genome  
database cDNA collection of Pioneer Hi-Bred International,  
Inc."  
1. 3538  
/gene="Cesa-6"  
43. 3222  
/gene="Cesa-6"  
/note="Cesa-6"  
/codon\_start=1  
/product="cellulose synthase-6"  
/protein\_id="AAF89966.1"  
/db\_xref="GI:9622884"  
/translation="MDORNGVQCICGDDVGRNPDEGPFVACNECAFPICRDCYEYER  
REGTONCPOCKTRFKRFGKCARVPGEDEEDVDLENEFNWSDKDSQYLAESMLHAH  
MSYGRADLDGVFPQFHPINPVPLITNGQWVDIPDOHALVPSFVGSGGKRIHPLPY  
ADPNLPVQPRSMDSKDLAYGYGVAVKERMESWKOKERMHQTRNDGGDDGDAD  
LPLMDARQPLSRKIPLPSSQINPYRMIIIRLVLCFFHYRVMHPVDAFALWLIS

BASE COUNT 876 a 813 c 928 g 920 t 1 others  
ORIGIN  
Query Match 64.6%; Score 2543; DB 8; Length 3538;  
Best Local Similarity 86.9%; Pred. No. 0;  
Matches 2821; Conservative 0; Mismatches 415; Indels 9; Gaps 2;  
OY 192 CTCGTCGTCATCCCGCCGCGACGCGGATCCCGGCGGAAGCCGCGCGGAGCAGAACGGG 251  
|||||  
DB 1 CTCGTCGTCATCCCGCCGCGATGCGGAGCCGGAAGCCCATGACCAACGGGAGCCG 60  
OY 252 CAGGTGTCAGATTTGCGGCGACGACGTCGCTTGCCCGCGCGGGAGCCCTTCGTG 311  
|||||  
DB 61 CAGGTGTCAGATTTGCGGCGACGACGTCGCGGCGCAACCCGACGCGGAGCCTTTCGTG 120  
OY 312 GCGTGCAACGAGTGCCTTCCCGCTGTCGCGGAGCTGCTACGAATACGAGCGCGGAG 371  
|||||  
DB 121 GCCTGCAACGAGTGCCTTCCCGCTGTCGCGGAGCTGCTACGAATACGAGCGCGGAG 180  
OY 372 GGCACGCAAGACTGCCCCAGTGCAAGACTGATTAAGCGCTCAAGGCGTCCAAAGT 431  
|||||  
DB 181 GGCACGCAAGACTGCCCCAGTGCAAGACTGATTAAGCGCTCAAGGCGTCCAAAGT 240  
OY 432 GTGACCGGTGACGAGAGGAGGAGCGGCGTGCATGACCTGACCAACGAGTTCAACTGG-- 488  
|||||  
DB 241 GTGACCGGTGACGAGAGGAGGAGCGGCGTGCATGACCTGACCAACGAGTTCAACTGGAGC 300  
OY 489 GACGGCCATGACTCGCAGTCTGTGCGCGCATGCTCTACGCGCCACATGACTACGGC 548  
|||||  
DB 301 GACCAAGCAGACTCCCAAGTACCTCGCGCATGCTCTACGCGCCACATGACTACGGC 360  
OY 549 CGTGGAGTGAACCTTAATGCGCGCCACAGCTTCCAGCTCAACCCCAATGTTCACTC 608  
|||||  
DB 361 CGCGGCGCGCACTTCAGCGCGCTGCGCGCATTCACCCCATCCCAATGTTCCCTC 420  
OY 609 CTCACCAACGCGCAATGCTGATGACATCCACCGGAGCAGACGCGCTGCTCT 668  
|||||  
DB 421 CTCACCAACGCGCAATGCTGATGACATCCCGCGGAGCAGACGCGCTGCTCT 480  
OY 669 TTCATGGGTGCGGGAAGAGATACATCCCTTCTTATGCGGATCCAGCTTACCT 728  
|||||  
DB 481 TTCATGGGTGCGGGAAGAGATACATCCCTTCTTATGCGGATCCAGCTTACCT 540  
OY 729 GTGCAACCAAGTCTATGACCCATCCAGAGATCTTGCTGCATATGCGTATGTT 788  
|||||  
DB 541 GTGCAACCAAGTCTATGACCCATCCAGAGATCTTGCTGCATATGCGTATGTT 600  
OY 789 GCTTGAAGAGACGATGAGAAATTGGAAGAGAGAGAGAGAGAGAGATGACACGAGG 848  
|||||  
DB 601 GCATGAAGAGAGATGAGAGCTGGAAGAGAGAGAGAGAGAGAGATGACACGAGG 660  
OY 849 AATGATGTTGCTGATGATGTCAGATGCTGATCTACCACTAATGATGAAGCAAGA 908  
|||||  
DB 661 AACGATGCGGCGCGATGATGTCAGATGTCAGATCTACCACTAATGATGAAGCTAGA 720  
OY 909 CAACAACGTCAGAGAAATTCACCTTCATCAAGCCAGATTAATCAATAGATGATT 968  
|||||  
DB 721 CAGCATGTCAGAGAAAGATCCCGCTTCTTCAAGCCAAATCAACCCCTATAGATGATT 780





Db 2941 AACTCATTTGAGTGTAGCTGCATCTCCATGCGATCAACAAGGATATGATCATGG 3000  
QY 3183 GGGCCCTCTTTGGGAGCTATTCTTGCATTTTGGGTGATGTCATCTTATCCCTT 3242  
Db 3001 GGGCCCTCTTTGGGAGCTATTCTTGCATTTTGGGTGATGTCATCTTATCCCTT 3060  
QY 3243 CTCAAAGTTTGGTTGGAAGCAAAACAGACACCAACGATTGTCATGCTGTGCATT 3302  
Db 3061 CTCAAGGTTCTGTTGGAGGACAGACGACCAACGATTGTCATGCTGTGCATT 3120  
QY 3303 CTGCTGCTCAATCTTCTCGCTCTTGGGTTGCGATTGATCCTTCTTCCGAAGAT 3362  
Db 3121 CTCTGCTTGCATCTTCTCGCTCTTGGGTTCCGATCGACCCGTTCTTCCGAAGAT 3180  
QY 3363 GATGTCGCTTCTTGGAGAGTGTGTTGATTCGAAGTATGATGATGATGATGATGAT 3422  
Db 3181 GATGTCGCTTCTTGGAGAGTGTGTTGATTCGAAGTATGATGATGATGATGATGAT 3240  
QY 3423 CCCCC 3427  
Db 3241 TCCCC 3245

RESULT 3  
AF200532 3812 bp mRNA linear PLN 31-AUG-2000  
LOCUS AF200532 Zea mays cellulose synthase-8 (Cesa-8) mRNA, complete cds.  
DEFINITION AF200532 Zea mays cellulose synthase-8 (Cesa-8) mRNA, complete cds.  
ACCESSION AF200532  
VERSION AF200532.1 GI:9622887  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 3812)  
AUTHORS Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S.,  
Xococonile-Cazares, B. and Delmer, D.P.  
TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene  
family  
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)  
MEDLINE 20398328  
PUBMED 10938350  
REFERENCE 2 (bases 1 to 3812)  
AUTHORS Dhugga, K.S. and Helentjaris, T.G.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred  
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA  
FEATURES  
source location/Qualifiers  
1..3812  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/note="identified and isolated from the maize genome  
database cDNA collection of Pioneer Hi-Bred International,  
Inc"  
1..3812  
/gene="Cesa-8"  
214..3498  
/gene="Cesa-8"  
/note="Cesa-8"  
/codon\_start=1  
/product="cellulose synthase-8"  
/protein\_id="AAF89968.1"  
/db\_xref="GI:9622888"  
/translation="MEASAGLVAGSHNRELIVIRDRREGAAGGAARRAEPQIC  
GDEVGVFDGEFVACNECAFVCAACYEREREGSQPCQTRKRLKGCPRVAG  
EEDGVDDLEGEFGLQDGAHEDDQYVAESMLRAQMSYRGDAPGFSPVNPVPL  
TNGWVDDIPPEOHALVPSYSGGGGKRIHPLPADNPVQPRSMDSKDLAAG  
YGSVWKEWMEGWKQOERLQVRSBGGDWDGDDADLPLMBARQPLSRKVPISSR  
INPYRMILIVIRLVVLGFFHYRVHPAKDAFALMLISVICELWFAMSWILDQPKMLP  
IERETYLDRLSLRFDEKQPSQALPIDFVSTVDPTKEPPLVANTVLSILSDYVE  
KVSCTVSDDGAAMLTPEALSETSEPAKKWVPFSKKFNIEPAPEMYFQCKIDYLDKDY

gene  
CDS

AAAFVRRERAMKREYEEKVRINALVAKAQKVPPEGTMDGSPWPGNNVRDHPGMIO  
VFLGSGGRDVEGENEDPLVYVSREKRPNGYNHKKAGAMNALVRVSAVLSNAAYLLNL  
DGDHYINNSKAIKEAMCFMMDPLVKKVCYVQFPOREFDIDKNDRYANRNVFEDIM  
KGLDGIQGIYVGTGCVFRQALYGDAPKTKPPSRTCQWPKLSCCSCRNKMK  
KTKPKTEKKRLPFKKAENPSPAYALGEIDEGAPGADIEKAGIVNOQKLEKFGQSS  
VEFASLLENGTILKSASPASLKEAIVHISGVEDKTDWKEIGWISITEDILTG  
FMHCHGWSRISYICIPKPAFKGSAPLNLSRLHQLVRLMALGSVEIFFSKHCPWYGYG  
GGLKFLERSYINSIYVWTSIPLAYCTLPALCITLTKFITPELTNVAISWFMALFI  
CISVTGILEMRWSGVAIDDMWRNEQFWVIGVSAHLFAVQGLKVPAGIDTSFTVTS  
KAGDDEBSELYTEFKMTLLIPTTLLNFIQVAVISNAINNGYESWGPFGKLPF  
AFWIVHLVPLFKGLVGRQNRPTPTIVIVWSILLASIFSLMWVRVDPFLAKSNGPLLEE  
CGLDGN"  
BASE COUNT 873 a 966 c 1111 g 862 t  
ORIGIN

Query Match 56.9%; Score 2241; DB 8; Length 3812;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;  
QY 123 GGAAGTGAAGGAGGAGGAAGCGATGAGAGCGCGCGGCTGTGTCGCGCTCCACAAC 182  
Db 193 GGGCGGGCGGGGGGCTGAGATGAGGCTAGCGCGGCTGTGTCGCGCTCCATAAC 252  
QY 183 CGCAAGAGCTCGTCATCCGCGCGAGCGGCGA-TCCCGGCGGAAGCGCGCGGGA 241  
Db 253 CGGAAGAGCTGTGTGATCCGCGCGAGCGGAGTCCGAGCGCGCGCGCGCGGCGG 312  
QY 242 GCAGACGGGCG----AGGTGTCCAGATTGGCGGCGAGCGTGGCTTGCCTCCGCG 296  
Db 313 GCGCGCGGGCGGAGGCGCGCGCTCCAGATATCGGCGAGCGAGTGGGCTTGCAC 372  
QY 297 GGGGACCCCTTCTGTGCGTGCACAGAGTCCGCTTCCCGCTGTCCGCGGAGTCTACGA 356  
Db 373 GGGGAGCCCTTCTGTGCGTGCACAGAGTCCGCTTCCCGCTGTCCGCGCTTACGAG 432  
QY 357 TACGAGCGCGGGAGGCGAGCGAGAGTCCCGGAGTGCAGAGTCTGATACAGCGCTC 416  
Db 433 TACGAGCGCGGAGGCGTGCAGAGCGTCCCGAGTGCAGAGCGCGCTACAGCGCTC 492  
QY 417 AAGGCTGCCAACGTGTGACCGGTGACGAGAGGAGGAGGAGCGCGTGCATGACCTGACAAC 476  
Db 493 AAGGCTGCCCGCGGCTGCGCGGCGAGAGGAGGAGGAGGAGCGCGTGCATGAGGCG 552  
QY 477 GAGTTCACTGG-----GACGCGCATGACTCGCAGTCTGTGCGCGAGTCC 521  
Db 553 GAGTTGGGCTTGCAGGACGCGCGCGCGCGAGAGGAGGAGCGCGCAGTACGTCGCGAGTCC 612  
QY 522 ATGCTTACGGCCACATGAGCTACGCGCGTGGAGGTGACCTTAATGCGCGCGCAACAAGCT 581  
Db 613 ATGCTACGGCGCAGATGAGTACGCGCGCGCGCGA-----CGCGACCCCGCGC 663  
QY 582 TTCCAGTCAACCCCAATGTTCCACTCTCTCAACCAACGCGCAATGTGATGACATCCCA 641  
Db 664 TTCAAGCCCGTCCCAACGTGCGCGCTCTCAACCAACGCGCAGATGTTGATGACATCCCG 723  
QY 642 CCGAGGAGCAGCGCGCTGTGCTTCTTATG-----GGTGTGGGGGAGAAAGAG 692  
Db 724 CCGAGGAGCAGCGCGCTGTGCGCGCTTACATGAGCGCGCGCGCGGGGCAAGAGG 783  
QY 693 ATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGTCTATGAGACCA 752  
Db 784 ATCCACCCGCTCTTTCGCAAGATCCCAACCTTCCAGTGCACCGAGATCCATGAGACCCG 843  
QY 753 TCCAAGATCTTGTCTCATATGGGTATGTGTGTTGTTGGAAGCAACGATGAGAAAT 812  
Db 844 TCCAAGATCTTGGCGCGCTTACGAGATATGCGAGCGGTGGCTTGAAGGAGAAATGAGGGC 903  
QY 813 TGAAGCAGAGCAAGAGAGATGACACGAGCGGGAATGATGTTGTTGTTGAT--GAT 869  
Db 904 TGAAGCAGAGCAAGAGAGCGCTGACAGATGTCAAGAGCGAGGTTGGTGTGAT 963  
QY 870 GGTGAGATGCTGATCTACCACTAATGATGAAGCAAGCAACCACTGTCCAGGAAAT 929

Db 964 GGGCAGATGCAAGATCTGCCACTAATGATGAAGCTAGGCAGCCAATTGTCCAGAAAACTC 1023  
QY 930 CCACTTCCATCAAGCCAGATTAACTCATATAGATGATGATATCATATATCCGCTTGTT 989  
Db 1024 CTTATATCATCAAGCCGAATTAACTCCCTACAGAGATGATATCTGTTATCCGGTTGGTGT 1083  
QY 990 TTGGGTTCTTCTTCCACTACCGAGTATGATCCGGTGAATGATGATGATGCTTGTGG 1049  
Db 1084 TTGGGTTCTTCTTCCACTACCGAGTATGATCCGGGAAAGATGATTTGCATTGTGG 1143  
QY 1050 CTCATATCTGTTATCTGTGAATCTGGTTGCCATGCTTGATCTTGATCTGATCAATCCCA 1109  
Db 1144 CTCATATCTGTTATCTGTGAATCTGGTTGCCATGCTTGATCTTGATCAATCCCA 1203  
QY 1110 AAGTGTCCCTATTGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTGACAAG 1169  
Db 1204 AAGTGGCTTCCAATCGAGAGAGACTTACCTGACCGTTGTCACTAAGGTTGACAAG 1263  
QY 1170 GAAGGCCAGCCATCTCACTTGTCCAATTGATTTCTTGTCACTACGGTTGATCCCTTA 1229  
Db 1264 GAAGGTCAACCCCTCTCAGCTTGTCCAATCGACTTCTTGTCACTACGGTTGATCCACA 1323  
QY 1230 AAGGAACCTCTTGTGTCAACAACAATACTGTTCTATCTATCTTTCGGTGATATCT 1289  
Db 1324 AAGGAACCTCCTTGTGTCAACAGCACTGCTTCCATCCTTCTGTGATTAATCCG 1383  
QY 1290 GTTGATTAAGTTTCTTGTATGTTTGTATGATGATGCTGCTCAATGCTTAACGTTGAAGCA 1349  
Db 1384 GTTGAGAAAGTCTCCTGTATGTTTGTATGATGATGCTGCTCAATGCTTAACGTTGAAGCA 1443  
QY 1350 TTATCTGAAAACTGAAATTTGCAAGAATGGGTTCCCTTTCGCAAAAGGTACAATATT 1409  
Db 1444 TTGTCTGAAAACTGGAATTTGCAAGAATGGGTTCCCTTTCAGCAAAAAGTTAATATC 1503  
QY 1410 GAACCTCGCGCTCCAGAGTGTACTTCCAACAGAAAGATAGACTTGAAGACAAAGGTG 1469  
Db 1504 GAGCCTCGTGTCTCTGAGTGTACTTCCAACAGAAAGATAGACTTGAAGACAAAGGT 1563  
QY 1470 GCAGCAAACTTGTAGGAGAGAGAGCAATGAAGAGAGATGAAGAAATCAAGGTG 1529  
Db 1564 GCTGCTTCAATTGTAGGAGAGAGAGCGCATGAAGAGAAATGAAGAAATCAAGGT 1623  
QY 1530 AGAATCAATGCTTAGTGGCAAAAGCCAGAAAGTTCCTGAAGAGAGATGACAATGCAA 1589  
Db 1624 AGGATCAATGCTTGTGTCAAAAGCCCAAAAGTTCCTGAAGAGAGATGACAATGCAA 1683  
QY 1590 GATGGAACCCCTGGCTGGAACAATGTTCTGTATCATCTGGAATGATTCAGTCTTC 1649  
Db 1684 GATGGAAGCCCTGGCTGGAACAATGTTCTGTATCATCTGGAATGATTCAGTATTC 1743  
QY 1650 CTTGGCCAAAAGCGAGGCTTGAAGTGAAGGAAATGAATGCACTGCCAGATTGTTATGTT 1709  
Db 1744 CTTGGCCAAAAGTGGCGGTCTGTATGGAAGGAAATGAATGCTCGCTGGTTATGTC 1803  
QY 1710 TCTAGAGAGAAAAGCAAGGCTATAACCATCATTAAGAAAGCTGGTCTATGAATGCAATG 1769  
Db 1804 TCGAGAGAAAAAGAGCCAGGTTATTAACCATCAAGAAGGCTGGTCCATGAATGCACTG 1863  
QY 1770 GTCCGAGTCTCTGCTACTAACAATGCTCCATATTGTTAACTTGATGATGATGATGAT 1829  
Db 1864 GTCCGTGTCTGTGTCTTATCAAAATGCTGCATACCTATGAACTTGAACTGTGATCAC 1923  
QY 1830 TACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTATGATGAGCCCTTACTA 1889  
Db 1924 TACATCAACAATAGCAAGGCCATAAAGAGGCTATGTGTTTCAATGATGATCCTTTGGTG 1983  
QY 1890 GGAAGAAGGTTTGCTATGTACAGTTCCTCAAAAGATTGATGGGATGATGCGCATGAC 1949  
Db 1984 GGGAGAAGAAAGTGTCTATGTACAGTTCCTCAGAGGTTGATGATTAAGCAAAAATGAT 2043  
QY 1950 CGATATGCTAACCGGAATGTTGCTTTTGTATATCAACATGAAGGTTGATGATGAT 2009  
Db 2044 CGATACGCTAACAGGAAGTGTCTTTTGTGACATCAACATGAAGGTTTGACGCTAT 2103

QY 2010 CAGGTCCAATTATGTTGTAATGATGATGATTTAGAAAGCAGGCAATTAATGCTTAT 2069  
Db 2104 CAAGACCATTATGTTGTAATGATGATGATTTAGAAAGCAGGCAATGATGCTTAT 2163  
QY 2070 GATGCCCCCAAAACAAGAAGCCACATCAAGACTTGCACTGTGGCCCAAGTGTC 2129  
Db 2164 GATGCTCTAAACAAGAGAGCCACATCAAGAACTTGCACTGTGGCCCAAGTGTC 2223  
QY 2130 TTTTGTCTGTCTGCTCTTGAGCAATAGGAAGCAAAAAGAGCTTACCAAAACCAAGAG 2189  
Db 2224 CTCTCTGCTGCTGCAAGCAAGAAATAAAAAGAGACTACAAAACCAAGAGCGAG 2283  
QY 2190 AAGAAAAAGTTATTTTCAAGAAAGAGAAACCAATCCCTGATATGCTCTTGT 2249  
Db 2284 AAGAGAAAAAGATTATTTTCAAGAAAGCAAAAAACCATCTCTGATATGCTTTGGT 2343  
QY 2250 GAAATGAAGAACTGCTCCAGAGCTGAGAAATGAAGAAAGCCGATATGTAATCAACAA 2309  
Db 2344 GAAATGATGAAGGTGCTCCAGGCTGATATCGAGAAAGCCGAAATCGTAATCAACAG 2403  
QY 2310 AAATTAAGAAAAAGAAATTTGGCCAACTCTGTTTGTATACATCAACTCTCGAAT 2369  
Db 2404 AAATTAAGAAAAAGAAATTTGGCCAACTCTGTTTGTATACATCAACTCTCGAAT 2463  
QY 2370 GGTGAACCTTGAAGAGTGCAGTCTGCTTCTCTTTGAAGAAAGCTATACATGTCAT 2429  
Db 2464 GGAAGGACCTGAAGAGCGCAAGTCCAGCTTCTCTGAAGAAAGCTATACATGTCAT 2523  
QY 2430 AGTTGCTTATGAAGACAGACAGACTGGGAAAAAGAGATTGCTGATCTATGATCA 2489  
Db 2524 AGCTGGCTACGAAGACAAAGACCACTGGGAAAAAGAGATTGCTGATTTACGATCG 2583  
QY 2490 GTTACAGAAGATATTCTAAGTGTTCAGATGATGATGATGATGATGATGATGATGAT 2549  
Db 2584 ATCAGAGAGATATCTGATGATTTAAGATGACACTGCCATGCTGGCGGTCTATTAC 2643  
QY 2550 TGCAATCCTAAACGGGTGCATTCAAAAGTTCTGCACCTCTGAATCTTCAAGTCTCT 2609  
Db 2644 TGCAATCCGAAGACGGCGCTGCATTCAAAAGTTCTGCACCTCTGAATCTTCAAGTCTCT 2703  
QY 2610 CACCAAGTCTTGGTGGCTCTTGGGCTATGAGATCTTCTGACCAATCATTCGCCCT 2669  
Db 2704 CACCAAGTCTTGGTGGCTCTTGGGCTATGAGATCTTCTGACCAATCATTCGCCCT 2763  
QY 2670 CTTTGGTATGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2729  
Db 2764 CTTTGGTATGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2823  
QY 2730 ATGCTATCTTGGATGACATCTATCCCTCTTGGCTTACCTGATGATGATGATGATGAT 2789  
Db 2824 ATGCTATCTTGGATGACATCTATCCCTCTTGGCTTACCTGATGATGATGATGATGAT 2883  
QY 2790 TTATGACAGGAAATTTATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGTTCATG 2849  
Db 2884 CTGCTCAAGGGAAGTTATCAACAGAGCTTACCAATGCTGCCAGTATCTGTTCATG 2943  
QY 2850 TCACTTTTATCTGCATTTTGTCTACAGCATCTTGAATGATGATGATGATGATGAT 2909  
Db 2944 GCACTTTTATCTGCATTTTGTCTACAGCATCTTGAATGATGATGATGATGATGAT 3003  
QY 2910 ATTGATGCTGTGAGGAATGACAGTCTTGGTCAATGAGGTGTCTCACACCTC 2969  
Db 3004 ATGACGACTGTGTGAGGAATGACAGTCTTGGTCAATGAGGTGTCTCACACCTC 3063  
QY 2970 TTTGCTGTGTCCAGGACTTCTCAAGTCAATGATGATGATGATGATGATGATGATGAT 3029  
Db 3064 TTTGCTGTGTCCAGGACTTCTCAAGTCAATGATGATGATGATGATGATGATGATGAT 3123  
QY 3030 ACATCAAGGCTGAGATGATGAGAGTTCTCAGAGCTATATACATTCAAATGACTACC 3089  
Db 3124 ACCTCAAGGCGCGGGAAGAGAGAGTCTCGAGCTGTACAGTTCAGTGAACCAACC 3183



QY	3090	TTATGTATACCTCCTACCACTTCTTCTTATTTGAACCTTATGTGTGTGCTGGCGTT	3149
Db	3184	CTGCTGATACCCCCGACCAAGCTCTCTCTGCTGAACCTTCAATCGGGGTGTGGCCGGATC	3243
QY	3150	TCAATGCGATCAATAACGGATATAGTCATGGGGCCCCCTTTGGGAAGCTATTCTT	3209
Db	3244	TGGAACGCGATCAACAACGGGTACGAGTCGTGGGGCCCCCTGTTGGGAAGCTCTTCTTC	3303
QY	3210	GCATTTTGGGTGATTTGCCATCTTTATCCCTTTCTCAAGCTTTGGTTGGAAGCAAAAC	3269
Db	3304	GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCTCCTCAAGGCTGTGGGAGGCGAAGAC	3363
QY	3270	AGGACACCAACGATTGTCATCGTCTGGTCCATCTGCTGCGCTTCAATCTTCTCGCTCTT	3329
Db	3364	AGGACGCGCAGATCGTCAATCGTCTGGTCCATCTGCTGCTGCTGATCTTCTCGCTCTG	3423
QY	3330	TGGCTTCGGATTGATCTTCTCTTCCGGAAGATGATGTCGCTTCTTGAGGAGTGTGT	3389
Db	3424	TGGTTCGCGCTGACCCGCTTCTCGCAAGACGACGCGCTCTGAGAGAGTGTGCGC	3483
QY	3390	TTGATTGCAACT	3402
Db	3484	CTGACTGCAACT	3496
RESULT 4			
LOCUS	AF304374	3847 bp	mRNA linear PLN 02-MAY-2001
DEFINITION	Nicotiana alata cellulose synthase catalytic subunit (CesA1) mRNA, complete cds.		
ACCESSION	AF304374		
VERSION	AF304374.1	GI:13925880	
KEYWORDS			
SOURCE	Nicotiana alata (Persian tobacco)		
ORGANISM	Nicotiana alata		
REFERENCE	Doblin, M.S., De Melis, L., Newbiggin, E., Bacic, A. and Read, S.M. Pollen tubes of Nicotiana alata express two genes from different beta-glucan synthase families Plant Physiol. 125 (4), 2040-2052 (2001)		
AUTHORS	Doblin, M.S., De Melis, L., Newbiggin, E., Bacic, A. and Read, S.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-SEP-2000) Plant Biology, University of California, 1 Shields Ave, Davis, CA 95616, USA		
FEATURES	location/Qualifiers		
source	1. .3847		
	/organism="Nicotiana alata"		
	/mol_type="mRNA"		
	/db_xref="taxon:4087"		
	/cissu_type="pistil"		
	/note="genotype: S2S2"		
	1. .3847		
	/gene="CesA1"		
	103. .3378		
	/gene="CesA1"		
	/codon_start=1		
	/product="cellulose synthase catalytic subunit"		
	/protein_id="AAK49454.1"		
	/db_xref="GI:13925881"		
	/translation="MDTKRLVAGSHNRNEFVINADDYGRVTSYKELSGQICQICGDD EIEVTVDGEPIACNECAPVCRQCYEYERREGNACPOCKTRFKIKGSPRYDGDDE DDEFDDLDEFDYHGNPRYMESEAFSSRLGRGTNNASGLTTPSEVDPAALNSEIPLLL TYGQEDDTISADKHALIIPFMGRGKXHPVPYSDMSLPPRPMDPKDLAVYGYGT AWKERMDWKKQNDKLOVVKHGGGKGNDGDELDDPLPKMDEGRQPLSRKLPIS SRSPYRLILVRLAVGLFFHYRITHPVNDAYALMLISICEIMEAVSNITDQFPKM PPIVRETYLDRLSLRYEKEKPSGLAPIDIFVSTVDPLKEPPLITANTVLSIAVDY EDKVSCTVSDGAAMLTFEALSETSEFARKWVPECKKNIEBPRAPEWYFSQKVDYLK		

Query Match	Best Local Similarity	Score	Pred. No.	Length	Matches	Conservative	Mismatches	Indels	Gaps
38.9%;	68.2%;	1531.4;	DB 8;	3847;	2252;	0;	996;	53;	7;
139	AAGCGATGAGGCGGCGCGCGGCTGTGGCGCGCTCCACCAACCGCAACGAGCTCGTCG	198							
98	AAGTAATGATACTAAAGGAGGCTAGTGTGTTTCAACAATAGGAATGAGTTGTG	157							
199	TCATCCGCGCGAGCGCGATCCCGGCGGAAGCCGCGGAGCAGAACGGGCGAGGT	258							
158	TCATCAATGCTGATGATGTTGGAAGAGTAACCTGTGAAGAATTAAGTGGCGAGATT	217							
259	GCCAGATTGCGCGCAGACGCTGCGCTTGCCCCCGGGGACCCCTTCTGTGCGTCA	318							
218	GCCAGATTGTGGAGATGAGATTGAAGTTACAGTAGATGGGAGCCATTATTGCTTCCA	277							
319	ACGAGTGGCGCTTCCCGCTCTGCCGGGACTGTCTACGAATACGAGCGCGGAGGGCACGC	378							
278	ACGAATGTGATTCCTGTTTGACAGCAATGCTTAGATATGAAGGAGAGAGAGAAATC	337							
379	AGAACTGCCCCAGTGCAGACTCGATACAAAGCGCTCAAGGGCTGCCAAGCTGTGACCG	438							
338	AAGCTTGCCCTCAATGCAGAACTCGTTCAAGCGCATTAAGGAGAGCCAGGTTGATG	397							
439	GTGACGAGAGAGAGACGGCGCTCGATGACCTGACCAACGAGTTCACTGGAGCGGCATG	498							
398	GAGATGATGAAGATGATGAAATTGATGACTTAGACCATGAGTTCCATTACCATGG-TAAC	456							
499	ACTGCGAGTGTGGCGGAGTCCATGCTTAC-----GGCCACATGAGCTACGGCGCT	551							
457	CCTGATATATGCTGAGGCTGCATTCTTCTGCTTGGCGCGGTACCAATCATATAT	516							
552	GGAGGTGACCCCTAATGGCGCGCCACAAG-----CTTCCAGCTCAACCCCAATGTTCCA	605							
517	GCTTCTGGGCTTACTACCCCATCAGAGTAGATCCTGCTCTTAACCTCAGAGATCCCT	576							
606	CTCCTCAACCAACGGGCAATGGTGGATGACATCCACCGAGACGACGCGCTGTGCT	665							
577	CTTCTTACTTATGCTCAAGAGGATGATACAAATTTCCGCTGACCAAGCATGCTTTATCAT	636							
666	TCTTTCATGGGTGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCAGCTTA	725							
637	CCACCGTTATGGGTGCGGGAAGAAAGTTCATCTGTGCTTATTCGATTCCA--TG	693							
726	CCTGTGCAACCCAGGTTATGAGACCCCATCAAGGATCTTCTGTCATATGGTATGTA	785							
694	TCTTTGCCCCCTGACCCCATGATCCTAAGAAAGATTTGGCAGTTTATGATATGGACT	753							
786	GTTGCTTGAAGAACCGATGGAGATTTGGAAGCAGACAGAGAGAGATGACCAACG	845							
754	GTTGCATGGAAGAAAGATGAGGAGCTGGAAGAAAAACAGAAATGATTAATTACAGGTG	813							
846	GGAATGATGCTGTGTGATGATGT-----GACGATGCTGAT	884							
814	GTTAAGCAGGAGCGGTAAAGGTGTGTTAAGATGAGATGAGCTGATGATCCGAT	873							
885	CTACCACTAATGATGAAGCAAGACAACTGTCCAGGAAAAATTCACCTTCATCAAGC	944							

QY	3090	TTATTGATACCTCCTACCACCTTGCTTCTATTGAAC TTCATTGGGTGGCGCCGTT	3149
Db	3184	CTGCTGATACCCCCGACCAGCCTCCTCCTGCTGAAC TTCATCGGGGTGGGCCGGGATC	3243
QY	3150	TCAAATGCATCAATAACGGATATGAGTCATGGGCCCTCTTGGGAAGCTATTC TT	3209
Db	3244	TCGAACGC GATCAACAACGGGTACGAGTCGTGGGCCCTGTTCGGGAAGCTTCTTC	3303
QY	3210	GCATTTTGGGTGATGTGCCATCTTATCCCTTCTCAAAGGTTGGTTGAAGGCAAA C	3269
Db	3304	GCCTTCTGGGTGATCGTCCAACCTGTACCCGTTCTCAAGGGTCTGGTGGAGGCA GAAC	3363
QY	3270	AGGACACCAACGATTTGTCATCGTCTGTGCATTTCTGCTGCTTCAATCTTCTGCT CCTT	3329
Db	3364	AGGACGCCGACGATCGTCATCGTCTGTCCATCTCTGCTGCGCTCGATCTTCTCG CTCTG	3423
QY	3330	TGGGTTCCGATTGATCTTCTTCTTGCGAAGGATGATGTC CGCTTCTTGAGGAGTGTG GT	3389
Db	3424	TGGGTCCGCGTCGACCCGTTTCTTCGCCAAGAGCAACGCGCCGCTCCTGAGGAG TGTGGC	3483
QY	3390	TTGGATTGCAACT 3402	
Db	3484	CTGACTGCAACT 3496	

RESULT 4  
AF304374

LOCUS	AF304374	3847 bp	mRNA	linear	PLN 02-MAY-2001
DEFINITION	Nicotiana glauca cellulose synthase catalytic subunit (CesA1) mRNA, complete cds.				

ACCESSION	AF304374
VERSION	AF304374.1
KEYWORDS	GI:13925880
SOURCE	Nicotiana glauca (Persian tobacco)

ORGANISM    *Nicotiana glauca*  
                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
                   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
                   Asteridae; Lamiales; Solanales; Solanaceae; *Nicotiana*.

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 3847)	Doblin, M.S., De Melis, L., Newbiggin, E., Bacic, A. and Read, S.M.	Pollen tubes of <i>Nicotiana glauca</i> express two genes from different beta-glucan synthase families	Plant Physiol. 125 (4), 2040-2052 (2001)

MEDLINE	21196092
PUBMED	11299383

REFERENCE 2 (bases 1 to 3847)  
AUTHORS Doblin,M.S., De Melis,L., Newbigin,E., Bacic,A. and Read,S.M.  
TITLE Direct Submission  
JOURNAL Submitted (12-SEP-2000) Plant Biology, University of California, 1  
Shields Ave, Davis, CA 95616, USA

FEATURES	Location/Qualifiers
source	1. .3847

```

/organism="Nicotiana glauca"
/mol_type="mRNA"
/db_xref="taxon:4087"
/tissue_type="pistil"
/note="genotype: S2S2"
1..3847

```

CDS	/gene="CesA1" 103.3378
-----	---------------------------

```

/gene="CesA1"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AK49454.1"
/db_xref="GI:13925881"
/translation="MDTKGRVLVAGSHNRNEFVYNADDVGRVTSVKELSGOICQIGDE
EIEVTVDGEPIACNECAFPVRCRCYERYEREGNACPQCKTRFKRIKGSPRVNDGDBB
DDEFDDLDFEFDYHGNNRMYSEAEFSSRLGRGTNNASGLTTPSEVDPAALNSEIPLLT
TYGQEDDTISADKHALIIPFMGRGKKVHPVPSDSMSLPPRMDPKDLAVYGYGTIV
AWKERMEDWKKKQNDKLQVVKHGSGKGKGGNDGDELDLDPDKMDEGRQPLSRKPLISS
SRLSPYRLLLVRLAVGLFHYRIHPVNDAYALWLISICEIWFVSWIFPDQPPKK
PIVRETYLDRLSRYEKEGKPSGLAPIDFVSTVDPLKEPRLTANTVLSILADVXE
EDKVSQVSDDGAAMLTEALSETSETSEFARKWVPECKSNIEPRAPEWYFSQKVLYLKN

```

Query Match	Similarity	Score	DB	Length
Best Local Match	68.2%;	Pred. No. 8.8e-286;		
Matches 2252;	Conservative 0;	Mismatches 996;	Indels 53;	Gaps 7;
BASE COUNT	1061 a	719 c	925 g	1142 t
ORIGIN				
Query	139	AAGCGATGAGGCGAGCGCGGGCTGTGCGCGCTCCACAACCGCAAGCAGCTCGTCG	198	
Db	98	AAGTAATGATACTAAAGGAGGCTAGTTGCTGGTTCAACAATAGAAATGAGTTGTTG	157	
QY	199	TCATCCGCGCGAGCGCGATCCCGGGCGAGCGCGCGGGAGCAGAACCGGCAGGTGT	258	
Db	158	TCATCAATGCTGATGATGTTGGAAAGATACTTCTGTGAAGAATTAAGTGGCAGATT	217	
QY	259	GCCAGATTTGCGGCGAGCAGCTCGGCTTGCCCCCGGGGAGACCCCTTCGTGGCGTGA	318	
Db	218	GCCAGATTTGTGAGATGAGATTGAAGTTACAGTAGATGGGAGCCATTATTGCTTGCA	277	
QY	319	ACGATGCGCTTCCCGCTGTCGGGAGCTGTACGAATACGAGCGCGGGAGGGCAGCG	378	
Db	278	ACGAATGTCATTCCTGTTTGACAGCAATGCTAGTAGATGAAGAAGAGAGAAATC	337	
QY	379	AGAACTGCCCCAGTGCAGAGCTGATACAAAGCGCTCAAGGCGTGCACGTGTGACCG	438	
Db	338	AAGCTTGCCCTCAATGCAAACTCGTTCAAGCGCATTAAGGAGCCCCAGGTTGATG	397	
QY	439	GTGACGAGAGGAGAGCGCGCTGATGACCTGGACAACGATTCAACTGGACGGCCATG	498	
Db	398	GAGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	456	
QY	499	ACTCGAGTCTGTGGCGGAGTCCATGCTCTAC-----GGCCACATGAGCTTACGGCCGT	551	
Db	457	CCTCGATATATGCTGAGGCTGCACTTCTTCTGCTGCTGCGCCGCGTACCAATCATAT	516	
QY	552	GGAGTGAACCTAATGGCGCGCCACAAG-----CTTTCCAGCTCAACCCCAATGTTCCA	605	
Db	517	GCTTCTGGGCTTACTACCCCATCAGAAGTAGATCTGCTGCTCTTAACCTCAGAGATCCCT	576	
QY	606	CTCCTCAACGAGGCAATGCTGATGACATCCCAACGAGCAGCAGCGCTGCTGCT	665	
Db	577	CTTCTTACTATGTCAGAGAGATGATACATTTCCGCTGACCAAGCATGCTTATCAT	636	
QY	666	TCTTTCATGGGTGTGGGGAAGAGATACATCCCTTCTTATGCGATCCAGCTTA	725	
Db	637	CCACCGTTATGGGTGCGCGAAGAAAGTTTATCTGCTGCTTATCTGATTTCA---TG	693	
QY	726	CCTGTGCAACCCAGGTCTATGACATCCCATCCAGAGATCTTGTGCTCATATGGGTAGT	785	
Db	694	TCTTTGCCCCCTCGACCCATGATCTTAAGAAAGATTGGCAGTTTATGATATGGGACT	753	
QY	786	GTTGCTTGAAAGAAAGGATGGAATTGGAAAGCAGAGCAAGAGAGATGACACAGACG	845	
Db	754	GTTGATGAGAGAAAGAAATGAGAGACTGGAAAGAAAACAGAAATGATTAATTAACAGTG	813	
QY	846	GGGAATGATGCTGTGTGATGATGCT-----GACGATGCTGAT	884	
Db	814	GTTAAGCAGAGGCGGTAAAGGTGTGTAAAGATGAGATGAGCTGATGATCCGAT	873	
QY	885	CTACCACTAATGATGAAGCAAGCAACACTGTCCAGGAAATTTCCACTTCCATCAAGC	944	

Db 874 TTGCCCAAGATGGATGAAGGTAGACAACCACTTCGAGAAAAGTCCTATTTCCTCTAGC 933

Qy 945 CAGATTAATCCATATAGGATGATTATCATTTATTCGGCTGTGGTTTGGGGTCTTCTTC 1004

Db 934 AGGCTAAGCCCATACAGATTGCTCATTTTAACTTCGTCCTGTGCTGGCTCTTTTTC 993

Qy 1005 CACTACCGAGTGTATGCATCCGGTGAATGATTCATTTGCTTTGGCTCATATCTGTATC 1064

Db 994 CACTATAGATAACGCATCCTGTCAATGATCATATGCGTGTGGCTCATTTCTATTATA 1053

Qy 1065 TGTGAATCTGGTTTGGCATGTCTTGGATTCTTGATCAATTTCCCAAAGTGTCCCTATT 1124

Db 1054 TGTGAATATAGTTTGGCGTATCATGGATATTGATCAGTTCCCAAATGGTTCCGATT 1113

Qy 1125 GAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTGACAAGAAAGGCCAGCCATCT 1184

Db 1114 GTGCGAGACATACCTGACAGGCTATCTGTAGGTATGAGAAAGAAAGGAAAGCCCTCT 1173

Qy 1185 CAACTTGCTCCAATTGATTTCTTTGTCAATACGGTTGATCCCTTAAAGAAACCTCCTTG 1244

Db 1174 GGGTTAGCCCCATAGATATATTGTTAGTACGGTGGATCCCTGAAGAACCTCCGCTT 1233

Qy 1245 GTCAACAATACTGTCTATCTATCCCTTCCGTGGATTATCCTGTGATAAGTTTCT 1304

Db 1234 ATTACTGCAATACTGTTCTCTCCATTCTCGCTGTGGATTATCCGAGACAAGGTTTCT 1293

Qy 1305 TGCTATGTTTCTGATGATGCTGCTCAATGCTAAAGTTGAAGCATTTATCTGAACATCT 1364

Db 1294 TGCTATGTTCTGACGATGCTGCTCCATGCTTACTTTGAGGCTCTCTCTGAAACATCT 1353

Qy 1365 GAATTTGCAAGAATAGGGTCCCTTCTGCAAAACGGTACAATATTGAACCTCGCGCTCCA 1424

Db 1354 GAGTTCGCAAGAAATGGGTCCCATCTGCAAGAAAGTTCAACATPAGACCTCGGCCCA 1413

Qy 1425 GAGTGTACTTCCAAAGAGATAGACTTGAAGAACAAGGTGCGACAAACTTGT 1484

Db 1414 GAGTGTATTCTCTCAGAAAGTTGATTATCTGAAAAATAAGTACATCCATCATTTGTG 1473

Qy 1485 AGGAGAGAGAGACAATGAAGAGAGATGAGAAATTCAGGTGAGAAATCAATGCCTTA 1544

Db 1474 AGGAAAGTCGTGCTATGAAGAGACTATGAAGAATTCAGGTTCCGATTAATGCACTG 1533

Qy 1545 GTTGCCAAAGCCAGAAAGTTCCTGAAGAAAGATGGACAATGCAAGATGGAAACCCCTGG 1604

Db 1534 GTTGCCAGGCACAAAAAGTTCCTGAAGATGGCTGGACCATGCAAGATGGAAACGCTTGG 1593

Qy 1605 CCTGAAACAATGTTGCTGATCATCTGGAATGATTCAAGTCTTCCGCAAAAGCGGA 1664

Db 1594 CCCGAAATCTTGTAGGATCATCTGGAATGATCCAGTCTTCCGTGGTAAATGATGT 1653

Qy 1665 GGCTTGACTGTGAGGGAATGAAGTGCACGATTGTTATGTTCTAGAGAAACGA 1724

Db 1654 GTCCGTATATTGAAGGAATGATGCTCGCCTTATCTATGTTCTCGTGAGAAGCGT 1713

Qy 1725 CCAGGCTATAACCATCATTAAGAAAGCTGTGCTATGAATGCATGCTCCGAGTCTGTCT 1784

Db 1714 CCAGGATTTGACCAACCAAGAAAGGCTGTGCCATGAATGCTTTGATGCGGGTATCAGCG 1773

Qy 1785 GTACTTAACAATGCTCCATATTTGTTAACTTGATGTGATCACTACATCAACAACAGC 1844

Db 1774 GTCATATCAAAATGCTCCTTACTTGCTCAATGTGATTGTATCAATATAATTAACAGT 1833

Qy 1845 AAGGCTATAAGGAACAATGTGTTTATGATGGACCTTTACTAGAAAGAGGTTTGC 1904

Db 1834 AAGGCACTGAGAGAAAGCTATGTCTTATGATGACCCCACTTCAGAGAAAGAAATATGC 1893

Qy 1905 TATGTACAGTCCCTCAAGATTTGATGGGATGATCGCCATGACGATATGCTAACCGG 1964

Db 1894 TATGTACAAATCCCAAAAGTTGATGATTTGACAGGATGACAGATACTCCAACCGC 1953

Qy 1965 AATGTTGCTTTTGTGATCAACATGAAGGTTTGATGTATTCAGGTCCAATTTAT 2024

Db 1954 AACGTGTCCTTCTTGATATCAATATGAAGAAAGACTTGACGGGATCCAGGTCCAATTTAC 2013

Qy 2025 GTTGTACTGATGTGTATTTAGAGCGACGCAATTTATATGTTATGATGCCCAAAACA 2084

Db 2014 GTGGTACTGATGTGTCTTTCAGAGAGGCCAAGCGCTTTATGTTATGATGCTCCAAAGAA 2073

Qy 2085 AAGAAGCCCATCAAGACTTTCGAACCTGCTGGCCCAAGTGTCTTTGCTGTGCTGC 2144

Db 2074 ACMAAACCTCCAGMAAAACATGCAACTGCTGGCCMAAATGTGTGCTGTGTTCCGT 2133

Qy 2145 TTGCGAATAGGAAGCAAAAGAGACTACCAAAACCAAAACAGAGAAAGAAAGTTATTA 2204

Db 2134 TCT-----CGAAAGAGCATTAAGAAAGGAAACAAACCAAGGATTAACAAAG 2181

Qy 2205 TTTTCAAGAAAGAGAAACCAATCCCTGCATATGCTCTTGTGAATTGACGAAGCT 2264

Db 2182 AAGACTAAACCAAGAGGCTTCAACCAAAATTCACGCCCTTGAATAATTCAGAGAAAG 2241

Qy 2265 GCTCCAGAGCTGAGATGMAAAGCCCGTATTGTAATCAACAAAATTAGAAAGAAA 2324

Db 2242 ATTGAAGGAATTGATAGTAAAGCAACCTCATGCTCAGATAAACTTGAGAAAGAA 2301

Qy 2325 TTGGCCAATCTTCTGTTTGTTCATCCACACTTCTCGAAGATGTGGAACCTTGAAG 2384

Db 2302 TTGGAACAATCACCAAGTATTCGTTGCTTCAACACTTCTAGAGATGTGTCATTCCTCA 2361

Qy 2385 AGTGCAGTCTGCTCTCTTTTGAAGAGCTATACATGTCATTAGTTGTGTTATGAA 2444

Db 2362 GGAGCAACATCAGCATCTCTCGTGAAGAAAGCAATACATGATAGTGTGTTATGAA 2421

Qy 2445 GACAGACAGACTGGGAAAGAGATTGGCTGATCTATGATCAAGTTACAGAAAGATTT 2504

Db 2422 GACAAACAGAAATGGGATAGAGAGTTGGATGATTTATGATCTGTTACTGAGGATATC 2481

Qy 2505 CTAACTGTTTCAAGATGCATGTCATGTTGGCGGTCAATTTACTGCATACCTAAACGG 2564

Db 2482 CTAAACGATTTAAGATGCATGTCACGGTTGAGATCTGCTAAGCATGCCAAAGG 2541

Qy 2565 GTTGCAATCAAAAGTTCGCACTCTGAATCTTTCAGATGCTTCAACAGTGTCCGG 2624

Db 2542 CTTGCTTTAAGGGGTCAAGTCTTATCAATCTTTCAGATGCTTTCATCAGGTTCTCGCG 2601

Qy 2625 TGGGCTTTGGGCTATGAGATCTTCTTCAGCAATCATGCCCTTTGGTATGGTAT 2684

Db 2602 TGGGCTTTGGGATCTGTTGAATTCCTTGAGTAAACATGTGCCATTTGGTATGATAT 2661

Qy 2685 GGTGGCGCTGGAATTTTGAAGAAAGTTTCTCTACATCAACTCCATCGTATCCTTGG 2744

Db 2662 GGGTGTGGGCTGAAGCCGCTGAGCGCTTTCCTACATAACTCAGTGTCTAACCCATTA 2721

Qy 2745 ACATCTATTTCCCTCTTGCTTACTGTACATTTGCCATGCTGTTTATGACAGGAAA 2804

Db 2722 ACTTCTCTCTTGTGATGCAATTTGTGCAATGCGAGCTGTGTTGCTTACTGGA 2781

Qy 2805 TTTATCACTCCAGAGCTGAATTAATGTTGCCAGCCTGTGTTCAATGCTTTTATCTGC 2864

Db 2782 TTTATGTTCCAGAGATTAGCAATATGCTAGCATTTTTCATGGGACTTTTCATAATG 2841

Qy 2865 ATTTTGTACGAGCATCTTGAATGAGATGAGTGTGTCTTCAACCTTTGTGTGTTCCAG 2924

Db 2842 ATTGCTGCCACGAGTCTCTGAGATGCAATGGGAGGTGTAATTTGATGACTGTGG 2901

Qy 2925 AGGAATGAGCAGTCTTGGGTATTTGAAGTGTGTCTTCAACCTTTGTGTGTTCCAG 2984

Db 2902 AGAAACGAGCAGTCTTGGGTATTTGGTGTGTCCTATCTCACTATTTGCTCTTTCCAA 2961

Qy 2985 GGACTTCTCAAGTCAATAGCTGTGTTGATACAAAGCTTCAACCGTGAACATCAAGGTTGA 3044

Db 2962 GGTCTTCTTAAAGTTTGGCTGTGTCAGCACTAGCTTCACTGTACTTCCAAAGCAGCA 3021

Qy 3045 GATGTAGAGAGTCTCAGAGCTATATACATTCAAAATGACTACCTTATGATACCTCT 3104

Db 3022 GACGACGGGAGTTTCTGAGCTTATCTGTTCAAGTGAACATCTTGTGATTCCTCT 3081



QY	3105	ACCACCTTGCTTCTAATGAACCTTCATTGGTGTGTCGCTGGCGTTTCAAAATGCGATCAAT	3164
Db	3082	ATGACTTTGTTAATTATTAACATCATTTGAGTCATAGTTGGAATTTCCGATGCAATCAAC	3141
QY	3165	AACGATATAGTCATGGGCCCCCTCTTTGGAGACATACTTTTGCAATTTGGGTGATT	3224
Db	3142	AACGGTATGATCTTGGGGCTCTATTCGGTAGCGTTTCTTCGCCCTGTGGGTCAAT	3201
QY	3225	GTCCATCTTTATCCCTTTCTCAAGGTTTGGTTGGAAGCAAAACAGACACCAACGATT	3284
Db	3202	GTTCATCTATATCCCTTCTCAAGGTGTGATGGGCAACAAATAAAGTTCCCAACCATC	3261
QY	3285	GTCATCGTGTGTCATTTCTGCTGGCTTCAATCTTCTCGCTCTTTGGGTTCCGATTGAT	3344
Db	3262	ATGTGTCTGTGCTATCTTCTTGCGCATCTATCTTTTCCTTAATGTGGGTCGAGTCAAC	3321
QY	3345	CCTTTCCTTGCGAAGGATGATGTCGCTTCTTGAGGAGTGTGTTGGATTGCAACTAG	3404
Db	3322	CCATTCACGGCAA--GAGGTGAACTCGTTTTAGAAGTATCGGGGTGGAACTGCGAGTAA	3378
QY	3405	G 3405	
Db	3379	G 3379	

RESULT	5
AX507835	
LOCUS	AX507835
DEFINITION	Sequence from Patent WO0216655.
ACCESSION	AX507835
VERSION	AX507835.1 GI:23389072
	3255 bp DNA linear PAT 27-SEP-2002

SOURCE ORGANISM	
Arabidopsis thaliana (chale cress)	
Arabidopsis thaliana	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	
1	

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Harper, J. F., Kreps, J., Wang, X. and Zhu, T.	Stress-regulated genes of plants, transgenic plants containing same, and methods of use	Patent: WO 0216655-A 2530 28-FEB-2002;

JOURNAL  
Patent: WO 0216655-A 2530 28-FEB-2002;  
The Scripps Research Institute (US) ; Syngenta Participations AG  
(CH)

FEATURES	Location/Qualifiers
source	1. .3255

```

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT      841 a      641 c      813 g      960 t
ORIGIN

```

Query Match	38.3%;	Score 1509;	DB 6;	Length 3255;
Best Local Similarity	67.6%;	Pred. No. 1.9e-281;		
Matches 2207; Conservative	0;	Mismatches 1030;	Indels 30;	Gaps 5.

QY 144 ATGAGGCGAGCGCCGGCTGTGCCCGCTCCCAACCGCAACGACTCCTCATC 203  
||| | | | | | | | | | | | | | | | |  
Db 1 ATGAACACCGGTGTCGGTTAATCGCCGGTCTCAACAAGAATGATTGCTCTCAT 60  
  
QY 204 CGCCGCACCGCGATCCCCGGGCCGAGCGCCCGGGAGCAGAACGGGACGTGTGCCAG 263  
||| | | | | | | | | | | | | | | | |  
Db 61 AATGCCGATGAGAATGCCCGAATAAGATCAGTCCAAGAGCTGAGTGACACATGTCAA 120  
  
QY 264 ATTTCGCGCGACGACGTGCGCTTGCCCCCGCGGGGACCCTTCGTGCCGTGCAACGAG 323  
||| | | | | | | | | | | | | | | | |  
Db 121 ATCTCACAGATGAGATCGAATTGACTGTGTGATGAGAAACCGTTTGTGCATGTAAACGAA 180  
  
QY 324 TGGGCTTCCCCGTCTGCCGGGACTGTCAAGATAAGAGCGCCGGAGGCAACGACGAAC 383  
||| | | | | | | | | | | | | | | | |  
Db 181 TTGTGATTTCCCTGTGTGTAGACCTTGCTATGAGTAAGAAAGACGAAAGGCATCAAGCT 240  
  
QY 384 TGCCCCCAGTGCAAGACTCGATACAAAGCGCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443

Db	241	TGTCACAGTGCAAAACCCGTTTCAACGCTTTAAAGAAAGTCCAGAGATTGAAGTGAT	300
QY	444	GAGGAGGAGGACGGCGTCGATGACCTGGACAAGAGTTCAACTGGGACGGCCATGACTCG	503
Db	301	GAAGAGGAAGATGACATGATGATTAGACAAATGAGTTTGA--GTATGAAATTAATGGG	357
QY	504	CAGTCTGTGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGAGGTGACCTT	563
Db	358	ATTGATTGTATCAGGTTTCTGAAGTATGTCAATCTCTCGCAACTCCGGTTTCCCA	417
QY	564	AATGGCGGCCACAGCCTTTCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA	623
Db	418	CAATCTGATTGGATTTCAGCTCCACCCTGGCTTCAGATTCCATTGCTGACTTACGGCGAC	477
QY	624	ATGCTGATGACATCCCAACCGGAGCAGCACCGCC--TGTGCTTCTTTCAATGGGTGT	680
Db	478	GAGGACGTTGAGATTCTTCTGATAGACATGCTCTTAATTGTTCTCTTCACTTGTGTGT	537
QY	681	GGGGAAAGAGAGATACATCCCTTCTTAAGCGATCCCACTTACCTGTCAACCCAGG	740
Db	538	CATGCAATAGAGTTCACTCTGTTCTCTTCTGACCCGACCGTGTGACATCCAGG	597
QY	741	TCTATGACCCATCCAAAGATCTTGTCTGATATGGGTATGTTAGTGTCTTGAAGAA	800
Db	598	CCTATGTACCTCAGAAAGATCTTCCGCTTATGTGTTATGGAAGTGTGCTGAAAGAT	657
QY	801	CGGATGGAATTTGGAAGCAGAGACAGAGAGAGATGCAACGACGGGA-----AT	851
Db	658	CGGATGGAAGATGGAAGAGAAAGCAGAAATGAGAACTTCAGGTTGTTAGGCATGAAGA	717
QY	852	GATGCTGTGTGATGATGCTGACGATGCTGATCTACCACTAATGATGAAGCAAGCAA	911
Db	718	GATCCTGATTTGAAGATGCTGATGATGCTGATTTCCAATGATGATGAGGAAGCAG	777
QY	912	CAACTGTCAGAAAAATTCACCTTCATCAAGCCAGATTATCCATATAGATGATTATC	971
Db	778	CCATTGTCTAGGAAGATTAACAATCAATCGAGCAAGATAAATCCTTACCGGATGTTAAT	837
QY	972	ATTATTCGGCTGTGGTTTTGGGGTCTTCTTCCACTACCGAGTGATGCATCCGGTGAAT	1031
Db	838	GTGCTACGCTTGATTTCTTGCTCTCTTCTTCACTACCGTATTTCTCAACCCCGTCAA	897
QY	1032	GATGCATTTGCTTTGGCTCATATCTGTTATCTGTGAATCTGTTGCCATGCTTTGG	1091
Db	898	GATGCATATGCTTTGGCTTATTTCTGTTATATGATGATATGTTGCTGTTCAATGG	957
QY	1092	ATTCCTGATCAATTCCTCAAGTGTTCCTTATGAGAGAGACTTACCTAGACCGCTG	1151
Db	958	GTTCTTGATCAGTTCCTTAATAGTATACCTATCGAGCAGAAACGTACTTGACCGACTC	1017
QY	1152	TCACTGAGGTTGCACAAGAACGCCAGCCATCTCAACTTGTCTCCAATTGATTCTTTGTC	1211
Db	1018	TCATTAAAGATATGAGAAAGAGGGAACCGTCGGGACTATCCCTGTGGATGATTGTGTT	1077
QY	1212	AGTACGGTGTATCCCTTAAAGGAACCTCCTTTGGTCAACAACAATACTGTTCTATCTATC	1271
Db	1078	AGTACAGTGTATCCATTGAAGAAGACCTCCGTTATTAATGCAAAATACGTCTGTCTATT	1137
QY	1272	CTTTCGGTGATTAATCCGTGTGATAAGGTTCTGTGCTATGTTTCTGATGATGCTGCTGA	1331
Db	1138	CTTGTCTGTGATTATCCCTGTGATTAAGGTTGCTGTTAGCTATCTGATGATGCTGCT	1197
QY	1332	ATGCTAACGTTGAAGCATTAATCTGAACATCTGAATTTGCAAGAAGATGGTTCCTTTC	1391
Db	1198	ATGCTTACTTTCGAAGCTCTTCTGAGACCGCTGAATTCGCAAGGAATGGGTTCCTTTC	1257
QY	1392	TGCAAAACGGTACAAATATTGAACCTCGCGCTCCAGAGTGTACTTCCAACAGAAATAGAC	1451
Db	1258	TGCAAGAAATATTGTATGAGCCTCGTGCTCCGAATGTATTCTGCCATTAATAATGAC	1317
QY	1452	TACTTGAAGACAAGGTGCGACGAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAG	1511

Db 1318 TACTTGAAGATAAAGTTCATCCCGCATTTGTTAGGGAGCGGCGAGCCATGAAGAGAGAT 1377  
QY 1512 TATGAGAATTCAAGGTGAGAAATCAATGCTTAGTTGCCAAAGCCGAGAAAGTTCTGAA 1571  
Db 1378 TATGAAGAATCAAGATAAGATCAATGCTTAGTAGCAACAGACAGAAAGTGCCTGAG 1437  
QY 1572 GAAGATGACAAATGCAAGATGGAACCCCGCTGGCCTGGAACAATGTTGATCATCCT 1631  
Db 1438 GATGTTGACTATGCAAGACGGTACACCTTGGCCCCGTAATAGTGTGAGATCATCCT 1497  
QY 1632 GGAATGATTCAGGTCTTCTTGCCAAAGCCGAGCCCTTGACTGTGAGGAAATGAACTG 1691  
Db 1498 GGCATGATTCAGGTCTTCTTGGAAGTGAACGGTGTTCGTGATGTCGAAAAACAAGATTG 1557  
QY 1692 CCACGATGGTTTATGTTTCTAGAGAGAAACGACGCTATTAACCATCAATAAGAAAGCT 1751  
Db 1558 CCTCGATTAGTTACGTTTCTCGTAGAAGAACCCGGATTGATCACCATAAAGAGGCT 1617  
QY 1752 GGTGCTATGAATGCATTTGTCGAGTCTGCTGTAATAACAATGCTCATATTGTTA 1811  
Db 1618 GGAGCTATGAATCCCTGATACGAGTCTGCGGTTCTATCAAAATGCTCCTTACCTCTG 1677  
QY 1812 AACTTGATGATCACTACATCAACAACAGCAAGGCTATAAAGAGCAATGTTT 1871  
Db 1678 AATGTCGATTGATCACTACATCAACAATAGCAAAAGCTTTAGAGAACAAATGTTTTC 1737  
QY 1872 ATGATGACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTTCCTCAAAAGATTGAT 1931  
Db 1738 ATGATGATCTCAGTCAGAAAGAAATCTGTTATGTTCAAGTTCCTCAAAAGTTGAT 1797  
QY 1932 GGGATTGATCGCATGACCGATATGCTAACCGGAATGTTGCTTTTGTATCAACATG 1991  
Db 1798 GGGATTGATGAGCAGATCGATCACTCAAAATGCAATGTTGTTGTTGATCAATATG 1857  
QY 1992 AAAGTTTGAATGATTCAGGTCCAATTATGTTGTAAGTACTGATGATTTAGAGG 2051  
Db 1858 AAAGTTTGAATGAGCTACAAGGCTATATACGTCGTAAGGTTGTTTTCAGAGAG 1917  
QY 2052 CAGCATTAATGCTTATGATGCCCCCAAAACAAGAGCCACCATCAAGACTTGCAAC 2111  
Db 1918 CAGCGCTTACCGATTGATGACCCGAGAGAGAGAGAGGCCACGTAAACATGCAAT 1977  
QY 2112 TGCTGGCCCAAGTGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGACT 2171  
Db 1978 TGCTGGCCCAATGCTGTCTCTATGTTTGGTTCAAGAAAGATCGTAAGCAAAAGACA 2037  
QY 2172 ACCAAACCCAAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAGAGAACCAATCC 2231  
Db 2038 GTGCTGCGATAGAGAGAGAGAA-----TAGGAAAGCGTCAAAAGCAGATC 2085  
QY 2232 CCTGCATATGCTTGTGGAATTGACGAAGCTGCTCCAGAGCTGAGATGAAAAAGGCC 2291  
Db 2086 CAGCATTAAGAAATATCGAAGAGGCGCGCTCACTAAAGTTCTAACGTAGAAGACATCA 2145  
QY 2292 GGTATTGAATCACAAAAATTAGAAAAGAAATTGGCCAAATCTCTGTTTGTGACA 2351  
Db 2146 ACCGAGGCAATGCAAAATGAAGTTGAGAGAAATTGGGCAGTCTCTGTATTGTTGCA 2205  
QY 2352 TCCACACTTCTCGAGAAATGCTGAACTTGAAGAGTCAAGTCTGCTTCTCTTTGAAA 2411  
Db 2206 TCTGCGCTATGAGAAATGCTGAGTGGCTAGAAACGCAAGCCCGCTGTCTGCTTAAA 2265  
QY 2412 GAAGCTATACATGTCATTAGTTGTGTTATGAGACAAGACAGACTGGGAAAAAGAGATT 2471  
Db 2266 GAAGCCATCCAAGTCATTAGTTGCGATATGAGATAAACTGAATGGGAAAAAGAGATT 2325  
QY 2472 GGCTGATCTATGATCAGTTACAGAAAGATTTCTAAGTGGTTTCAAGATGATGTCAT 2531  
Db 2326 GGGTGAATCTATGCTTCTGTTACCGAAGATATCTTACGGGTTTAAGATGATCTCAT 2385  
QY 2532 GGTGCGCGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAGGTTCTGCACCTCTG 2591  
Db 2386 GGTGAGATCTGTTATTGTACACCAAGATTAGCGGCTTTCAAGAGATCAGCTCCAATC 2445

QY 2592 AATCTTCAGATCGTCTTCAACGAGTGCTTCGGTGGGCTCTTGGGTCTATTGAGATCTTC 2651  
Db 2446 AATCTTCGATCGTCTCCATCAAGTTCTTCGATGGCGCTTGGGTGAGATTTTC 2505  
QY 2652 TTCAGCAATCATGCCCCCTTTGGTATGGGTATGTTGGCGGTCTGAAATTTTGGAAAGA 2711  
Db 2506 TTGAGTAGGCAATGTCCTATTGGTATGGTTATGAGAGGTGGGTGAATGCTTGAGCGG 2565  
QY 2712 TTTTCTACATCAACTCCATCGTATCTTGACATCTATTCCTCTTGCTTACTGT 2771  
Db 2566 TTGTCTTACATTAACCTGTGTTTAAACCGTGAGACCTCTTACCGCTCATCGTTACTGT 2625  
QY 2772 ACATTCCTGCCATCTGTTTATGACAGGAAATTTATCACTCCAGAGCTGAATAATGTT 2831  
Db 2626 TCTCTCCCTGCCATCTGTCTTCACTGGAATTTATCATCGTCCGAGATTAGCAACTAT 2685  
QY 2832 GCCAGCTGTGCTCATGTCACCTTTTATCTGCATTTTGTCTACGACATCCTAGAAATG 2891  
Db 2686 GCGAGTATCTCTTCAATGCGCTCTTCTCGTCAATTGCAATTAACGGGTATTCTCGAGATG 2745  
QY 2892 AGATGAGTGTGTTGGAATTGATGACTGTGAGAGAAATGACAGTCTGGTCATTGGA 2951  
Db 2746 CAATGGGCAAAAGTTGGATCGATGATGTTGGTAGAAGAACGAAAGTTTGGTCATTGGA 2805  
QY 2952 GGTGTCTCTCACACCTCTTTGCTGTGTTCCAGGACTTCTCAAGGTCATAGCTGTGTT 3011  
Db 2806 GGTGTCTTGCAGCATCTGTTGCTCTTCCAAAGTCTCCTCAAGTCTTGTGCTGTGTC 2865  
QY 3012 GATACAGCTTCACCGTGACATCAAGGCTGAGATGATGAGAGTTCTCAGAGCTATAT 3071  
Db 2866 GACACTAATTTCAAGTCAACATCAAAAGAGCTGATGATGAGAGATTCTGACCTTTAC 2925  
QY 3072 ACATTCAAATGGAATCACTATGATACCTCTTACCACTTGTCTTATTGAATTCATT 3131  
Db 2926 CTCTTCAATGGAATCACTTCACTTCACTTCAATGATCTCACTCATATAAGCTCAT 2985  
QY 3132 GGTGTGCTGCTGCGGTTTCAATGCGATCAATTAACGATATGATGAGGCCCCCTC 3191  
Db 2986 GAGTCAATGTCGAGTCTGTATGCGATCAGCAATGATGATGACTCGTGGGAGCCGCT 3045  
QY 3192 TTTGGAAGCTATTCTTTCATTTGGGTGATTGTCCATCTTATCCCTTCTCAAGGT 3251  
Db 3046 TTGGAAGACTGTTCTTTCACCTTGGGTGATCATTCATCTTAAACGTTCTTAAAGGT 3105  
QY 3252 TTGTTGGAAGCAAAACAGAGACCAAGATTGTCATGCTGTTCCATTCTGCTGCT 3311  
Db 3106 TTGCTTGGGAAACAAGATAGATGCCAACCATTTATGTCGTGCTGCCATCTCTGCGC 3165  
QY 3312 TCAATCTTCTGCTCTCTTGGGTGCAATGATCCTTCTTCCGGAAGGATGATGTCG 3371  
Db 3166 TCGATTCTTACACTTCTTGGGTGCGGTTAATCCGTTGTGCGA---AAGCGGTCT 3222  
QY 3372 CTCTTGAAGAGTGTGTTGATGTC 3398  
Db 3223 ATTCTGAGATCTGTGTTAGACTGC 3249

RESULT 6  
AF062485 3444 bp mRNA linear PLN 08-AUG-1998  
LOCUS AF062485  
DEFINITION Arabidopsis thaliana cellulose synthase mRNA, partial cds.  
ACCESSION AF062485  
VERSION AF062485.1 GI:3135610  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; euroids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 3444)  
AUTHORS Wu, L., Joshi, C.P. and Chiang, V.L.  
TITLE Aracela, a new member of the cellulose synthase gene family from

Arabidopsis (Accession No. AF062485) (PGR98-114)  
Plant Physiol. 117 (3), 1125 (1998)  
REFERENCE  
2 (bases 1 to 3444)  
AUTHORS  
Joshi, C.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (01-MAY-1998) Plant Biotechnology Research Center, School  
of Forestry and Wood Products, Michigan Technological University,  
Houghton, MI 49931, USA

FEATURES  
Source  
Location/Qualifiers  
1..3444  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/clone\_11b="ABRC Number CD4-16"  
<1..3247  
/codon\_start=2  
/product="cellulose synthase"  
/protein\_id="AAC29067.1"  
/db\_xref="GI:3135611"  
/translation="RPRLIAGSHNRNEFLINADENARIRSVOELSGTCOICRDEIE  
LTVDEPFVACNECAFVPCRCYEREREGNOACPQCKTRFKRLKGSPRVEGDEEDD  
IDDLNFEYGNNGTGFDOVSEGMISIRNSGFQSDLDAPSGOIPLITYGDEDE  
ISSDRHALIVPSPSLGGHNRVHVPVSLDPTVAHRRLMVPQKDLAVYGVSVAMKDRM  
EEMKRONKQLQVVRHEDDPDEDADDFPMWDEGRQPLSMKIPKSKINPYRMLI  
VLRLVILGLFPHYRIHPVKDAYALMLISVCEIMFASVWLDPKWPYERETYL  
RLSLRYEKEKPSGLSPDVVSTVDPLKEPLITANTVLSILAVDYPDKVACYSD  
DGAAMLTFEALSETAFARKWPFCKYCIERABEWYFCHMDYLKNKHPAFVRER  
RAMRDYEEFKKINALVATAOKVPEDEGTMODGTPWPGNSVRDHPMIQVFLGSDV  
RDVENNELPRLVYVSREKRBFEDHKKAGAMNSLIRVSGVLNAPYLNWCDHYINN  
SKALREAMCFMMDPOSEKKICYQVOPQRFQDIDRDRYSNRNVFFDINKGLDGLQG  
PIVGTGCVFRQALYGFDAKKKKRPKTCNCMPKMLLCFGSRKNRKAKTVAADKK  
KNREASKQIHLENIIEGRKVLNVEQSTEAMQMLQKYGQSPVFVASARLENG  
MARNASPACLLKEAIVISRGYEDKTEWKEIGNTIGSVTEIDILTSKMSHGWHRVY  
CTPKLAFFKGSAPINLSDRLHQLRWALGSVEIFLSRHCPIMYGGGLKWLRLSYI  
NSVVPWTSPLIVGSLPAICLLTGKFIVPKISNVASILFMALFSSIAITILEMOM  
GKVGIDDMWRNEQFWVIGVSAHLFALFOGLKVLAVDTNFTVTSKAADGEFSDLY  
LFKWTSLIIPMTLLINIVIGIVGSDAISNGYDSWGPLFGRLLFALWVILHLYPFL  
KGLLGKQDRMPTIIVWSILLASILTLLWVRVNPFAKGPILICGLDCL"

BASE COUNT 903 a 664 c 860 g 1017 t  
ORIGIN

Query Match 37.9%; Score 1489.8; DB 8; Length 3444;  
Best Local Similarity 67.3%; Pred. No. 9.8e-278;  
Matches 2189; Conservative 0; Mismatches 1032; Indels 30; Gaps 5;

QY 160 GGCTGGTGGCCGGCTCCACAAACGCAAGAGCTGCTCATCCGCGGAGCGCGATC 219  
DB 9 GGTAAATCCCGGTTCTCACAAACGAATGAGTTGTCTCATTAATGCCGATGAAATG 68  
QY 220 CCGGCGGAAGCCGCGGAGAGCAAGGCGAGGTGTGCGAGATTGCGGCGAGCAGC 279  
DB 69 CCGAATAAGATCATGTCGAAGAGCTGAGACAGACATGTCAATCTGCAGAGATGAGA 128  
QY 280 TCGGCTTGGCCCCGGGGGAGACCCCTTCGTGGCGTGCAACGAGTGCCTTCCCGCTC 339  
DB 129 TCGAATTGACTGTGATGAGAACCGTTTGTGCGATGTAAACGATGTGCAATCCCTGTGT 188  
QY 340 GCCGGAGCTCTACGAATACGAGCCCGGAGAGGCGACGACGAACTGCCCCAGTGCAGA 399  
DB 189 GTAGACCTTGCTATGAGTACGAAGACGAGAGGCAATCAAGCTTGTCCACAGTGCAAAA 248  
QY 400 CTCGATACAGCGCCTCAAGGCGTCCCAACGTGTGACCGGTGACGAGAGAGAGAGCGCG 459  
DB 249 CCGGTTCAACGCTCTTAAAGGAAGTCCCAAGAGTTGAAGGTGATGAGAGAGAGAGATGACA 308  
QY 460 TCGATGACTGACAAAGAGTTCAACTGAGAGCGCCATGACTCGAGTCTGTGCGCGAGT 519  
DB 309 TTGATGATTTAGACAATGAGTTTGA---GTATGAAATTAATGGATTGGATTGATCAGG 365  
QY 520 CCATGCTTAAGGCGCACATGAGTACGCGCGTGAGGTGACCTTAATGCGCGCGCAACAG 579  
DB 366 TTTCTGAAGGATATGTCATCTCTGTCGCACTCCGGTTTCCCACAATCTGATTTGATTT 425

QY 580 CTTTCAGCTCAACCCCAATGTTCACATCTCTCACCACGCGCAATGCTGATGACATCC 639  
DB 426 CAGCTCCACCTGGCTCTCAGATTTCCATTGCTGCACTTACGCGCAGCAGACGTTGAGATTT 485  
QY 640 CACCGAGCAGCAGCGGC--TGTGCTTCTTTCATGGGTGTGGGGGAAAGAGATAC 696  
DB 486 CTTCTGATAGACATGCTCTTATTGTTCTCTTCACTTGGTGTGATGCAATAGAGTTTC 545  
QY 697 ATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCAGGTCTATGACCCATCCA 756  
DB 546 ATCCTGTTCTCTTCTGACCCGACCGGTGCTGACATCGAAGGCTGATGACTTACAG 605  
QY 757 AGATCTTGCTGATATGGGTATGAGTGTGCTTGAAGAAAGAGATGAGAAATTGGA 816  
DB 606 AAGATCTTGGGTTATGTTATGATGAGAGTGTGCTTGAAGATCGATGAGAGATGGA 665  
QY 817 AGCAGACACAAGAGATGACACCGAGCGGGA-----ATGATGCTGTGTGATG 867  
DB 666 AGAGAAAGCAGATGAGAAACTTCAGGTTGTTAGGCATGAAGAGATCTGATTTTGAAG 725  
QY 868 ATGCTGAGATGCTGATCTACCACTAATGATGAGAACAGACAACTGCTCCAGGAAA 927  
DB 726 ATGCTGATGATGCTGATTTTCCAAATGATGATGAGGGAAGGAGCAGCATGCTATGAGA 785  
QY 928 TTCACCTTCATCAAGCCAGATTATTCATATAGATGATGATATCATATTCGGCTGTGG 987  
DB 786 TACCAATCAATCGACCAAGATAAATCTTACCGGATGTTAATGTGTACGTCCTTGTGA 845  
QY 988 TTTTGGGTTCTTCTTCCACTACCGAGTGAATGATCCGGTGAATGATGATGCTTGTGT 1047  
DB 846 TTCTTGCTCTTCTTCACTACCGTATCTTCAACCCCGTCAAGATGATGCTTGTGT 905  
QY 1048 GGCTCATATCTGTATCTGTGAATCTGTTGCCATGCTTGGATCTTCATCAATTC 1107  
DB 906 GGCTTATTTCTGTATATGTGAGATATGTTGCTGTTTCATGAGTTTCATCAAGTTCC 965  
QY 1108 CAAAGTGTTCCTTATGAGAGAGAGACTTACCTAGACCCGCTGTCACTGAGGTTGACA 1167  
DB 966 CTAATGTGATCCCTATCGAGCGAGAAACGTACTTGAACCGACTCTCATTAAGATATGAGA 1025  
QY 1168 AGGAAGGCCAGCATCTCAACTGTGCTCCAAATGATTTCTTGTCAATACGCTGATCCCT 1227  
DB 1026 AAGAAAGGAACCGTCGGAGACTATCCCTGTGATGATTTGTTAGTACAGTGAATCCAT 1085  
QY 1228 TAAAGGAACCTCTTGTGTCACAACAATACTGTTCTATCTATCCCTTGGTGAATATC 1287  
DB 1086 TGAAGAGCCTCCGCTATTACTGCAATACTGTGTCTATCTTGTGCTGATGATATC 1145  
QY 1288 CTGTTGATAAGTTTCTGCTATGTTCTGATGATGCTGCTGCAATGCTAACGTTGAAG 1347  
DB 1146 CTGTCGATAAGTTGCTTGTACGTATCTGATGATGCTGCTGCTTACTTTCGAAG 1205  
QY 1348 CATTAATGAAACATCTGAATTGCAAGAAATGGGTTCTTCTGCAACCGTACAATA 1407  
DB 1206 CTCTTTCTGAGACCGCTGAATTCGCAAGAAATGGGTTCTTCTGCAAGAAATATTGTA 1265  
QY 1408 TTGAACCTCGCGCTCCAGAGTGTACTTCCAAAGAGATAGACTACTTGAAGACAGAG 1467  
DB 1266 TTGAGCCTGCTGCCGATGATTTCTGCAATAATGACTACTTGAAGAAATGAAG 1325  
QY 1468 TGGCAGCAAACTTGTGAGGAGAGAGAGCAATGAGAGATGAGGAATTCAGG 1527  
DB 1326 TTCAATCCCGCATTTGTTAGGGAAGCGGAGCCATGAGAGATTAAGAAATTCAGG 1385  
QY 1528 TGAGAAATCAATGCTTGTGCAAGAGCCAGAAAGTTCTGAAAGAGATGAGCAATGC 1587  
DB 1386 TAAAGATCAATGCTTTAGAGCAACAGACAGAAAGTGCCTGAGATGTTGAGACTATGC 1445  
QY 1588 AAGATGAACCCCTGGCTGGAAGCAATGTTGATCATCTGGAATGATTCAGGTCT 1647  
DB 1446 AAGACGGTACACCTTGGCCCGTAATAGTGTGCGAGATCATCTGCAATGATTCAGGTCT 1505



Qy	1648	TCCTTGGCCMAAGCGGAGCCCTTGACTGTGAGGAAA	TGAATGAACTGCCACGATTGGTTATG	1707
Db	1506	TCCTTGAAGTGACGGTGTTCGTGATGTCGAAAAC	ACGATTGCCCTCGATTAGTTACG	1565
Qy	1708	TTTCTAGAGAAACGACCAGGCTATAACCATCAT	AAGAAAGCTGGTCTATGATGCAT	1767
Db	1566	TTTCTCGTAGAAGAGACC	CGGATTGTGATCACCATAGAAAGCTGGAGCTATG	1625
Qy	1768	TGGTCCGAGTCTCTGCTGTACTAACAAATGCT	CCATATTGTTAACTTGGATTGTATC	1827
Db	1626	TGATACGAGTCTCTGGGGTCTATCAATGCTCT	TACTCTCGAATGTCGATTGTATC	1685
Qy	1828	ACTACATCAACAACAGACAGGCTATAAAGGA	AGCAATGTTTTATGATGACCTTTAC	1887
Db	1686	ACTACATCAACAATAGCAAGCTCTTAGAAG	CAATGTTTCATGATGATCCTCAGT	1745
Qy	1888	TAGAAAGAAGGTTTGCATGTACAGTTCCCT	CAAGATTGATGGATTGATGCCATG	1947
Db	1746	CAGAAAGAATACTGTATGTTCAGTTCCCT	CAAGGTTGATGGATTGATGGCAG	1805
Qy	1948	ACCGATATGCTAACCCGAATGTGTCTTTT	TGATATCAATGAAGGTTTGATGTA	2007
Db	1806	ATCGATACTCAATCGCAATGTGTCTTTG	ATATCAATGAAGGTTTGATGGC	1865
Qy	2008	TTACAGGCTCCAATTATGTGTGTACTGTG	ATGTATTTAGAAAGCAGGCATTATATG	2067
Db	1866	TACAAGGGCCTATATACGTTCGGTACAG	GTGTGTTTCAGAGGCAAGCGCTTAC	1925
Qy	2068	ATGATGCCCCCAAAACAAGAACCA	CCATCAAGACTTGCAACTGCTGGCC	2127
Db	1926	TTGATGCACCGAAGAGAGAGAGG	CCACGTAAGACATGCAATTGCTGG	1985
Qy	2128	GCTTTTGCCTGTGCTGCTTTGSCA	TAGGAAGCAAAAGACTAACCAACC	2187
Db	1986	GTCCTCATATGTTTGGTTCAGAAAG	AATCGTAAAGCAAGACAGTGGCTG	2045
Qy	2188	AGAAAGAAAAGTTATTATTTTCA	AGAAAGAGAACCAATCCCTGCATAT	2247
Db	2046	AGAAAGAAA-----TAGGAA	GCCTCAAGAGATCCACGATTAGAA	2093
Qy	2248	GTGAAATTGACGAAGCTGCTCCAG	AGCTGAGATGAAAGCCGGTATTG	2307
Db	2094	TCCGAAAGGGCGCGCTCATTA	AGTTCTTAACGTAGAACAGTCA	2153
Qy	2308	AAAAATTAGAAAAGAAATTTG	CCCAATCTTCTGTTTTGTATCAT	2367
Db	2154	TGAAGTTGCAGAAAGAAATATG	GGCAGTCTCTGTATTGTTCAT	2213
Qy	2368	ATGCTGAACCTTGAAGAGTGA	AGTCTCTCTTTGAAAAGAGCT	2427
Db	2214	ATGCTGGGATGCTAGAAACG	CAAGCCCGCTGTCTGCTTAA	2273
Qy	2428	TTAGTTGTGTTATGAAGAC	AGACAGACTGGGAAAAGATTG	2487
Db	2274	TTAGTCGGGATATGAAGATA	AAACTGAAATGGGGAAGAGAT	2333
Qy	2488	CAGTTACAGAAAGATATTCT	TAACCTGTTTCAAGATGCAT	2547
Db	2334	CTGTTACCGAAGATATTCTTA	CGGGTCTAAGATGCATTCT	2393
Qy	2548	ACTGCATACCTAAACGGG	TTGCATCAAAAGTTCTGCAC	2607
Db	2394	ATTGTACACCAAAAGTTAG	CGGCTTCAAAAGATCAG	2453
Qy	2608	TTACACAGGTGCTTCGGT	GGCTTTGGGTCTATTGAGAT	2667
Db	2454	TCCATCAAGTTCTTCGAT	GGCGCTGGGTGCGTTGAG	2513
Qy	2668	CTCTTTGGTATGGGTATG	TGCGGCTCGAAATTTTGA	2727
Db	2514	CTATTGGTATGTTATG	AGGTGGGTTGAAATGCT	2573
Qy	2728	CCATCGTGTATCCTTG	ACATCTATTCCTCTTGCTT	2787

Db	2574	CTGTGTTTACCCGGGAGACCTCTCTACCGCTCATCGTTTACTGTCTCTCCCTGCCATCT	2633
QY	2788	GTTTATTGACAGGAAATTATCACTCCAGAGCTGAATATATGTTGCCAGCCTGTGTTCA	2847
Db	2634	GTCCTCTCACTGGAAATTCATCATCGTTCGCCAGATTAGCAACTATGCGAGTATCCTTCA	2693
QY	2848	TGTCACCTTTTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGAGTGTGTTG	2907
Db	2694	TGGCGCTCTCTCGTCGATTGCAATTAACGGGTATCTCGAGATGCAATGGGCAAAAGTTG	2753
QY	2908	GAATTGATGACTGGTGAGGAATGAGCAGTTCTGGGTCATTGGAGGTGTCTCACACC	2967
Db	2754	GGATCGATGATTGGTGAGAAACGAACAGTTTGGGTCATTGGAGGTGTCTCGCCATC	2813
QY	2968	TCTTGTCTGTGTCCAGGGACTTCTCAAGGTCATAGCTGTGTGATACAAAGCTTCAACC	3027
Db	2814	TGTTTGTCTCTTCCAAAGGTCTCTCAAGGTTCTGTGTGTGTCGACACATTAATTCAACAG	2873
QY	3028	TGACATCAAAAGGCTGAGATGATGAGAGTTCTCAGAGCTATATACATTCAAATGGAATA	3087
Db	2874	TCACATCAAAAGCAGCTGATGATGAGAGTTCTTGACCTTTACCTTCAAAATGGAATT	2933
QY	3088	CCTTATTGATACCTCCCTACCACTTGTCTTCTTAATGAATTCAATTGATGTGTGCTGGCG	3147
Db	2934	CACCTTCTCATCCCTCCATGAGTACTCTACTCATCATTAACGTCATTGGAGTCATGTCGGAG	2993
QY	3148	TTTCAATGCGATCAATAACGGATATGAGTCATGCGGCCCCCTCTTGGGAAGCTATTCT	3207
Db	2994	TCTCTGATGCCATCAGCAATGATAGTACGACTCGTGGGGAACCGCTTTCGGAAGCTGTTCT	3053
QY	3208	TTCATTTTGGGTCATGTCATCTTATCCCTTCTCAAAAGTTTGGTTGGAAGCAAA	3267
Db	3054	TTGACATTGGGTCATCATTCATCTTACCCGTTCTTAAAGTTTGGTTGGAAGCAAA	3113
QY	3268	ACAGAGACCAACGATGTCATCGTCTGTCGATCTGCTGCTTCTCAATCTTCTGCTCC	3327
Db	3114	ATAGAATGCCAACCATTAATTGTCTGCTGTCATCTCTGCGCTCGATTCTTACACTTC	3173
QY	3328	TTTGGGTTCCGATTCATCTTCTCTGCGAAGATGATGTCGCGCTTCTTGAGAGTGTG	3387
Db	3174	TTTGGGTCGGGTTAATCCGTTTGTGGCGA--AAGCGGCTCTATTCTCGAGATCTGTG	3230
QY	3388	GTTTGATTCG 3398	
Db	3231	GTTTAGACTGC 3241	
RESULT 7			
LOCUS	AF027173	3771 bp	mRNA linear
DEFINITION	Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-A)		
ACCESSION	AF027173		
VERSION	AF027173.1	GI:2827140	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J. and Williamson, R.E.		
TITLE	Molecular analysis of cellulose biosynthesis in Arabidopsis		
JOURNAL	Science 279 (5351), 717-720 (1998)		
MEDLINE	98111412		
PUBMED	9445479		
REFERENCE	2 (bases 1 to 3771)		
AUTHORS	Arioli, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-SEP-1997)		Plant Science Centre, Australian National

FEATURES		University, Acton, Canberra, ACT 200, Australia
Source	Location/Qualifiers	
	1..3771	
	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
	/cultivar="Columbia"	
gene	/db_xref="taxon:3702"	
	1..3771	
	/gene="Ach-A"	
CDS	217..3471	
	/gene="Ach-A"	
	/note="RSM1-like"	
	/codon_start=1	
	/product="cellulose synthase catalytic subunit"	
	/protein_id="AAC39335.1"	
	/db_xref="GI:2827141"	
	/translation="MNTGRLIAGSHNREFVLINADESARIRSVQELSGOTCOICGD EIELTVSSELEFVACNECAFVPCRPCEYERREGNOACPQCKTRYKIKGSPRVDDDE EEDIDILEYEFDHGMDPEHAEAALSRNTGRGLDSAPPGSQIPLITYCDEDADM YSDRHALIVPSTGYGNRVYPAPFTDSSAPFQARSMVPQKDLAEYGYGSVAMKDMEV WKRRQGEKLOVIKHGEGNNNGRSDNDELDDPMDPMDEGRQPLSRKLPISRINPY RMLILCRLAILGLFEHYRILHPVNDAYGLMTSVICEIWFVAVSWILDQEPKWPPIERE TYLDRLSLRYEKEGKPSGLAPVDVFSVTDPLKEPPLITANTVLSILAVDYPVDKVA YVSDDAAMLTFEALSDTAEFARKWVPCKFENIEPRAPEWYFSQKMDYLKNKHVPAF VREBRAMKRDYEEFKVKNALVATAQKVEEGWTMDGTWPFGNNVVDHPGMIQVFLG HSGVRDIDGNEPLRLVYVSREKPGFDHKKAGANSLIRSAVSNAPYLINVDCH YINNSKAIRESCFMMDPQSGKVCYVQFPQFDDIGDRHRYSNRVVFPDINMGDL GIOGPIYVGTGCVFRQALYGFDAPKKKPKPKTCNCMPKWCCLCCGLRKSKTKAKD KKTNTKETSKQIHLENVDEGVIIVPSNVEKRSEATQKLEKKFGSPVFVASAVLQN GGVPRNASPACILREAIQVISCYEDKTEWKEIGMIYGSVTEIDLTFKMHCHWRS VYCMPIKRAAFKGSAPINLSDRLHQLVRLALGSVEIFLSRHCPWYGGGLKWLERS YINSVVPWTSPLPLIVYCSLPVAVCLTGKFIIVEISNYAGILFMLEFISIAVTGILEM OMGVGIDDMWRNEQFVWIGASVSLFALFOGLKVLGAVNTNFTVTSKAADDGAFSE LYIFKWTLLIPPTLLIINIIGVIVGSDAISNGYDSWGPLFGRLEFALMVIIVHLYP FLKMLGKQDKMPTIIVVMSILASILTLMVRVNPFAKGGPVLKICGLNCGN"	
BASE COUNT	1037 a 699 c 907 g 1128 t	
ORIGIN		
Query Match	36.4%; Score 1434.2; DB 8; Length 3771;	
Best Local Similarity	65.9%; Pred. No. 5.7e-267;	
Matches 2175; Conservative	0; Mismatches 1063; Indels 63; Gaps 4;	
QY	124 GAAGTGAGGGAGGAGGAGATGAGGCGGCGGCGTGTGCGCGCTCCCAAC	183
Db	197 GATTCGGGTAGAGACATCATGATACTGTGTGCGCTCATGTGCTCTCAACA	256
QY	184 GCAACGAGCTCGTGTATCCGCGGACGCGGATCCCGGCGGAGCGCGGGAGC	243
Db	257 GAAACGAATTCGTTCTCAITTAACGCCGATGAGAGTGCCAGATAATCGATCA	316
QY	244 AGAAGGGGAGGTGTGCGAGATTGCGGCGAGACGTCGCTTGCCCCCGGGGACC	303
Db	317 TGAATGGGCAACATGTCAATCTGTGAGATGAATCGAATTAACGTTAGCAGT	376
QY	304 CCTTCGTGGCGTGAACGAGTGGCGCTTCCCGCTGCGCGGAGTGTACGATA	363
Db	377 TCTTTGTGCTTGCAACGATGCGCATTTCCCGGTTGTAGACCATGTCTAT	436
QY	364 GCGCGGAGGCGACGCAAACTGCCCGGAGTGCAAGACTCGATCAAGCGCT	423
Db	437 GTAGAGAAAGAAATCAAGCTTGTCTCAGTGCAAAACTCGATACAAAAGAT	496
QY	424 GCCAAGCTGTGACCGGTGACGAGGAGGAGGACGCGGTGATGACCTGACA	483
Db	497 GTCCACGGGTGATGAGATGAGAAGAAAGAAAGACATGATGATCTTGAT	556
QY	484 ACTGGAGCGGCATGATCGAGTCTGTGCGGAGTCCATGCTTACGGCCACA	543
Db	557 TTGATCATGGGATGACCTGAAACATGCGCGTGAAGCGCACTCTTCA	616
QY	544 ACGGCGGTGAGGTGACCTTAATGCGCGCCACAAGCTTTCAGCTCAAC	603
Db	617 CCGGTGCTGTGATGATGATTCAGCTCCACCTGCTCT-----CAGATT	661

QY	604 CACTCTCACCAACGGGCAATGTGTGATGATCCACCAGGACGACGCGGTGTC	663
Db	662 CTCTTTTGACTTAITGTGATGAAGATGCTGATATGTTCTGATCGTCATGCT	721
QY	664 CTCTTTTCAATGGGTGTGGGGAAAGAGATACATCCCCCTTCTTATGCGGAT	723
Db	722 TGCCCTCTCAACGGGATATGGAAATCGCGTCTATCCTGACCGGTTACAGAT	781
QY	724 TACCTGTGAACCCAGGTCTATGACCATCCAAAGATCTTGCTGATATGGTAT	783
Db	782 CACCTCCACAGCGGAGATCAATGCTTCTCAGAAAGATATGCGGAATATGCT	841
QY	784 GTGTGCTTGAAGGAACCGATGAGAAATTTGAAGCAGAGACAGAGAGATGCA	843
Db	842 GTGTGCTTGAAGGACCGTATGGAAGTTTGAAGAGACGACAGCGAAAGCTT	901
QY	844 CGGGGAATGATGCTGTGTGATGATGCT-----GACG	876
Db	902 TCATTAGCATGAAGAGGAAACAATGGTTCAGGTTCCAATGATGACGAACT	961
QY	877 ATGCTGATCAACCACTAATGATGAAGCAACAACAACCTGTCAGAAATTTCC	936
Db	962 ATCCTGACATGCTTATGATGATGAAGGAAGAACCACTCTCTCAGAAAGCT	1021
QY	937 CATCAAGCCAGATTAATCCATATAGATGATTAATATTGCGCTGTGTTGGG	996
Db	1022 GTTCAAGCAATAAATCCTTACAGAGATGTAATCTGTGTGCGCTCGGAT	1081
QY	997 TCTCTTCACTACCGAGTGATGATCCGGTGAATGATGATGATGCTTGTGCT	1056
Db	1082 TTTCTTCAITTAAGAAATCTCCATCCAGTCAATGATGATGATTAATGTTA	1141
QY	1057 CTGTATCTGTGAATCTGGTTTGCCATGCTTGATTTCTTGATCAATTC	1116
Db	1142 CAGTTATATGAGATATGTTGTGAGTGTCTTGATTTGATCAATTC	1201
QY	1117 TCCCTATTGAGAGAGACTTACCTAGACCGGCTGCTGAGTTCGACAAAGG	1176
Db	1202 ATCTATAGACGTGAACATACCTGATAGACTCTCTCAGATGACGAAAG	1261
QY	1177 AGCCATCTCACTTGCTCCAATGATTTCTTGTCACTACGTTGATCCCTTA	1236
Db	1262 AACCGTCAGATTAGCACCTGTGATGTTTGTAGTACAGTGAATCCGTTGA	1321
QY	1237 CTCTTTGTGACAAACAATACTGTCTATCTATCTCTTCCGTGATATCTGT	1296
Db	1322 CACCTTGATTAACGAAACACAGTCTTCCATTTAGCAGTGAATATCTGTG	1381
QY	1297 AGTTTCTGTATGTTTCTGATGATGCTGCAATGCTAACGTTGAAGCAT	1356
Db	1382 AGTTGCGTGTATGATACAGACGATGCTGACCTATGCTTACATTTGAAG	1441
QY	1357 AAACATCTGAATTTGCAAAAGAAATGGTCTTCTGCAAAACGTAACAAT	1416
Db	1442 ATACAGCTGAGTTGTAGAAAATGGTCTTTTGTGAAGATTAATATCGAG	1501
QY	1417 GCGCTCAGAGTGTACTTCCAAACAGAGATGACTACTTGAAGCAAGGTG	1476
Db	1502 GAGCTCCTGAGTGTATTTTCTCAGAGATGATTACCTGAAGCAAAAGTT	1561
QY	1477 ACTTGTAGGAGAGAGAGCAATGAGAGAGATATGAGAAATCAAGGTGA	1536
Db	1562 CTTTGTCAAGGAAAGTGTGCTATGAAGAGATTAATGAGAGATTAAAGTA	1621
QY	1537 ATGCTTAGTTGCCAAAGCCAGAAAGTTCTGAAGAGATGACAAATGCAAG	1596
Db	1622 ATGCACTGTTGCTACTGACAGAAAGTCCGAGGAAGTGGACTATGCAAG	1681
QY	1597 CCCCCTGGCTGGAACAATGTTCTGTATCATCTGGAATGATTCAGGTCTT	1656
Db	1682 CTCCTTGCTGGAACAACAGTCCGTGACCATCTGGAATGATTCAGGTCTT	1741





ORIGIN

Query Match		36.4%;	Score 1431.8;	DB 6;	Length 3255;
Best Local Similarity		66.0%;	Pred. No. 1.6e-266;		
Matches 2166;		Conservative	0;	Mismatches 1052;	Indels 63; Gaps 4;
QY	144	ATGAGGCGAGCGCGCGCTGGTGGCCGCTCCCAACCGCAACGAGCTCGTCATC	203		
Db	1	ATGAATACTGCTGGCTCATGTCTGCTCTCAACAGAAACGAATTCGTTCTCATTT	60		
QY	204	CGCCCGCAGCGCGATCCCGGCGCGAAGCCGCGCGAGCAGAACGGCGAGTGTCCAG	263		
Db	61	AACGCCGATGAGAGTGCAGAAATACGATCAGTACAGAACTGAGTGGGCAACATGTCAA	120		
QY	264	ATTGCGGCGACGACGTCGGCCTTGCCCCGCGGGACCCCTTCGTGGCGTGCAACGAG	323		
Db	121	ATCTGTGAGATGAATCGAATTAAACGTTAGCAGTGAAGCTCTTTGTGCTTGCAACGAA	180		
QY	324	TGCGCCTTCCCGCTCGCGGGAAGCTGCTACGAATACGAGCGCGGAGGCGACGACAAC	383		
Db	181	TGCGCATTCGCGGTTGTAGACCATGTCTATGAGTATGACGTAGAGAGAAATCAAGCT	240		
QY	384	TGCCCCCAGTGCAGACTCGATACAAAGCGCCTCAAGGCGTCCCAACGTGTGACCGGTGAC	443		
Db	241	TGTCCTCAGTGCAAAACCTCGATACAAAAGGATTAAAGTAGTCCACGGGTGTGATGAGAT	300		
QY	444	GAGGAGGAGGACGGCGGTGATGACCTGGACAACGATTCAACTGGAGCGGCCATGACTCG	503		
Db	301	GATGAAGAGAGAAAGACATTGATGATCTTGAGTATGAGTTTGATCATGGATGACCCCT	360		
QY	504	CAGTCTGTGGCGGATGCTCTACGCGGCACATGAGCTACGGCGGTGAGGTGACCT	563		
Db	361	GAACATGCCGCTGAAGCCGCACTCTCTTACGCGCTTAAACACCGGTCGTGTGATTGGAT	420		
QY	564	AATGGCGCGCACAAAGTTTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA	623		
Db	421	TCAGCTCCACCTG-----GCTCTCAGATTCCCTCTTTGACTTATTGTGAT	465		
QY	624	ATGCTGATGACATCCCAACGAGACGACGCGCTGCTTCTTTCATGGGTGGG	683		
Db	466	GAAATGCTGATATGTATTTCTGATCGTCATGCTCTTATCGTCCCTCTCAACGGGATAT	525		
QY	684	GGAAGAGGATACATCCCTTCTTATGCGGATCCGAGCTTACCTGTGCAACCGAGTCT	743		
Db	526	GGAATCGCTCTATCTCGACCGTTTACAGATTCTTCTGCACTCCACAGGCGAGATCA	585		
QY	744	ATGGAACCATCCAAGATCTGCTGATATGGTATGTTAGTGTGCTTGGAAAGAACGG	803		
Db	586	ATGTTCTCTCAGAAAGATTTGCGGAATATGTTATGGAAGTGTGCTTGGAAAGACCGT	645		
QY	804	ATGGAATTTGAAGCAGACAAAGAGATGACACGAGCGGGAATGATGTGTGT	863		
Db	646	ATGGAAGTTTGAAGAGACGACAGCGGAAAGCTTCAAGTCAATTAAGCATGAAGAGGA	705		
QY	864	GATGATGCT-----GACGATGCTGATCTACCACTAATG	896		
Db	706	AACAATGCTCGAGTTCCAAATGATGACGACGAACTAGATCTTGACATGCCCTATGATG	765		
QY	897	GATGAAGCAAGACAACACTGCTCCAGAAATTTCCACTTCCATCAAGCCAGATTAATCCA	956		
Db	766	GATGAAGAGAGACAACCTCTCTCAAGAAAGCTACCTATTCGTTCAAGCAGAAATAATCT	825		
QY	957	TATAGGATGATTATCATATTCGGCTGTGTGTTTGGGTTCTTCTTCCACTACCGAGTG	1016		
Db	826	TACAGAGATGTTAATTCGTGTGCGCTCGCGATTCTTGCTTTCTTTCATTATAGAAATT	885		
QY	1017	ATGCATCCGCGTGAATGATGATTGCTTGTGCTCATATCTGTATCTGTGAATCTGG	1076		
Db	886	CTCCATCCAGTCAATGATGATGATGATTATGTTAAGTCAGTATATATGCGAGATATGG	945		
QY	1077	TTTGCCATGTCTTGATTTCTGATCAATTCACCAAGTGTTCCTATTTGAGAGAGACT	1136		
Db	946	TTTGAGTGTCTTGATTTCTTGATCAATTTCCCAATGGTATCTATAGAACGTGAACA	1005		

QY	1137	TACCTAGACCGGCTGTCACTGAGGTTGCAAGGAAGGCCAGCCATCTCAACTTGCTCCA	1196		
Db	1006	TACCTCGATAGACTCTCTCTCAGGTACGAGAAAGAAAGAAACCGTCAAGATTAGCACCT	1065		
QY	1197	ATTGATTTCTTTGTCACTACGGTTGATCCCTTAAAGGAACCTCTTTGGTCAACAACAAAT	1256		
Db	1066	GTTGATGTTTGTGTAGTACAGTGGATCCGTTGAAAGAGCCACCCTTGATTACAGCAAAAC	1125		
QY	1257	ACTGTTCTATCTATCCCTTTCGGTGATTAATCCTGTTGATTAAGGTTTCTTGCTATGTTCT	1316		
Db	1126	ACAGTTCTTTCATTTCTAGCAGTTGATTAATCCTGTGATTAAGGTTGCGTTATGTATCA	1185		
QY	1317	GATGATGCTGCTCAATGCTAAAGTTGAAGCATTATCTGAACATCTGAATTTGCAAG	1376		
Db	1186	GACGATGGTGACGTATGCTTACATTGAAAGCTCTCTGTGATACAGCTGAGTTTGCTAGA	1245		
QY	1377	AAATGGTTCCTTTCTGCAACGGTACAAATTTGAACCTCGCGCTCCAGAGTGTACTTC	1436		
Db	1246	AAATGGTTCCTTTTGTAAAGATTATATTCGAGCCACGAGCTCTGAGTGTATTTT	1305		
QY	1437	CAACAGAAATAGACTACTTGAAGAAGAGGTGGACGAAACTTTGTTAGGAGAGAGA	1496		
Db	1306	TCTCAGAAATGGAATTAACCTGAAGAACAAAGTTTATCTCTGCTTTTGTCAAGGAACGTCT	1365		
QY	1497	GCAATGAAGAGAGATGATGAGAAATTCAGGTGAGAAATCAATGCTTAGTGCCAAAGCC	1556		
Db	1366	GCTATGAAGAGAGATTATGAGAGGTTAAAGTGAAGATAAATGCACTGCTGTCTACTGCA	1425		
QY	1557	CAGAAAGTTCCTGAAGAAGATGACATGCCAAGATGGAACCCCTGCGCTGGAACAAT	1616		
Db	1426	CAGAAAGTCCCTGAGGAAGGTTGGACTATGCCAAGATGGAACCTCTTGCGCTGGAACAAC	1485		
QY	1617	GTTGATGATCATCTCGAATGATTCAGGCTTCTCTTGCCAAAGCGGAGGCTTGACTGT	1676		
Db	1486	GTCGATGACCATCTCGAATGATTGAGGTGTTCTTGCGTCATAGTGAAGTTCGTGATACG	1545		
QY	1677	GAGGAATGAACTGCCAGATTGTTATGTTCTAGAGAGAAACGACGAGCTATTAAC	1736		
Db	1546	GATGTAATGAGTTTACCACGCTAGTGTATGTTTCTCGTAGAAGCGGCTTGATTTGAT	1605		
QY	1737	CATCATAGAAGCTGCTGCTATGAATGATGCTCCGAGTCTGCTGTACTAACAAT	1796		
Db	1606	CACCACAAGAAAGCTGAGATGATGAATTCCTTGATCCGAGTCTGCTGTCTATCAAAC	1665		
QY	1797	GCTCATATTTGTTAAACTTGATTTGATTCACCTACATCAACAACAGCAAGGCTATAAG	1856		
Db	1666	GCTCCTTACCTTCTTAATGTGATGTGATCACTACATCAACAACAGCAAAAGCAATTAGA	1725		
QY	1857	GAAAGATGTTTTTATGATGAGCCCTTACTAGGAAAGAAAGTTTGCTATGTACAGTTC	1916		
Db	1726	GAACTATGTTTTTATGATGAGCCCGCAATCGGGAAAGAAAGTTTGTATGTTCAGTTT	1785		
QY	1917	CCTCAAGATTGATGGGATGATGCGCATGACCGATATGCTAACCGGAATGTTGCTTT	1976		
Db	1786	CCGCAAGATTGATGGGATGATGACATGATAGATTAATCAACCGTAACGTTGTGTTT	1845		
QY	1977	TTTGATATCAACATGAAGGTTTGATGCTATTCAAGGTCCAATTTATGTTGTTACTGGA	2036		
Db	1846	TTTGATATTAACATGAAGAGTCTTGATGGGATACAGAGCCGATATATGTCGGGACAGGT	1905		
QY	2037	TGTGATTTAGAGCGCAGGATATATGTTATGATGCCCCCAAAACAAGAACCCACCA	2096		
Db	1906	TGTGTTTGAAGACAGGCTCTTATGCTTTTGATGCAACCAAGAAAGAAACCAACCA	1965		
QY	2097	TCAAGACTTGCACTGCTGGCCCAAGTGTCTTTTGTGTTGCTGCTTTGGCAATAGG	2156		
Db	1966	GGCAAAACCTGTAACTGTTGGCCTTAAATGCTGTTGTTGTGTTGGGTTGAAGAAAGAG	2025		
QY	2157	AAGCAAAAGAGACTAACCAACCCAAAACAGAGAAAGAAAGTTTATTTTTCAGAAA	2216		
Db	2026	AGTAAACGAAGCCAAAGATGAAGAAACTTAACACTTAAAGAGACTTCA-----	2073		

Qy	2217	GAAAGAAACCAATCCCTCGATATGCTCTTGTGAAATTGA	CGAAGCTGCTCCAGAGCT	2276
Db	2074	-----AAGCAGATTCA	TGCGCTAGAGAATGTCGACGAAGGTGTTATCGTCCCA	GTGCA 2127
Qy	2277	GAGATGAAAAGGCCGGTATTGTAATCA	CAAAAATTAGAAAAGAAATTTGGCCAATCT	2336
Db	2128	AATGTTGAGAAGAGATCTGAAGCA	CAACAACTTGAATTTGAGAAGAAAGTTGGACAATCT	2187
Qy	2337	TCTGTTTGTGTACATCCA	CTCTCGAATGGTGAACCTTGAAAGTGC	AAGTCT 2396
Db	2188	CCGGTTTCGTTGCTCTGCTGTCTTAC	AGAACGGTGAAGTTC	CCCGTAACGCAAGCCCC 2247
Qy	2397	GCTTCTCTTTGAAAGAGCTATAC	TGTCATTAGTTGTGTTATGAAGACAAGACAGAC	2456
Db	2248	GCATGTTTGTTAAGAGAAGCATTC	CAAGTTATTAGCTGGGGTACGAAGATAAAACCGAA	2307
Qy	2457	TGGGAAAAGAGATTGGCTGATCT	ATGATCAGTTACAGAAGATATTTCTA	CTGCTTC 2516
Db	2308	TGGGAAAAGAGATCGGGTGA	TTTATGATCGGTGATCAAGATATCTGACGGGTTTC	2367
Qy	2517	AAGATGCATTGTCA	TGGTGGCGGTCAATTACTGCATACCTAAACGGGTGCATTCAAA	2576
Db	2368	AAGATGCATTGCCATGATGAGATCT	GTGTACTGTATGCTTAAGCGTGACGCTTTTAAA	2427
Qy	2577	GGTCTGCACCTCTGAATCTT	CAGATCGTCTTACACAGTGCTTCGGTGGCTCTGGG	2636
Db	2428	GGATCTGCTCTATTAACTTG	TCAGATCGTCTTCACTAAGTTCTACGTTGGGCTCTGGC	2487
Qy	2637	TCTATTGAGATCTTCTTCA	GCAATCATTTGCCCTTTGGTATGGGTATGTTGGCGGTCTG	2696
Db	2488	TCTGTAGAGATTTTCTTG	AGCAGACATGTCCGATATGATGTTATGTTGTGTTTA	2547
Qy	2697	AAATTTTGGAAAAGATTTCT	CATCATCACTCCATCGTGTATCCTTGACATCTATTTCC	2756
Db	2548	AAATGTTGGAGAGATTCTCT	TACATCAACTCTGTCTATCTTGACTTCACTTCCA	2607
Qy	2757	CTCTTGCTTACTGTACAT	TGCTGCCATCTGTTATTGACAGGAAATTTATCACTCCA	2816
Db	2608	TTGATCGTCTATTGTTCT	CTCCCGCGGTTTGTACTACAGAAAATTCATCGTCCCT	2667
Qy	2817	GAGCTGAATATGTTGCC	AGCCTGTGTTCATGTCACTTTTATCTGCATTTTGGCTACG	2876
Db	2668	GAGATTAAGCACTACCG	CAGTATACTCTTCATGTCTCATGTTCATATCCATAGCAGTAACT	2727
Qy	2877	AGCATCTTAGAAATGAGAT	GAGTGTGTTGAAATGATGACTGTGGAGGAATGACAG	2936
Db	2728	GGAATCCTCGAAATGCAAT	GAGGAGGTGCGGAATCGATGATTGGTGAGAAAACGACAG	2787
Qy	2937	TTCCTGGCTATTGAGGT	GTGTCTCAACCTCTTGTGTGTCTCAGGACTTCTCAAG	2996
Db	2788	TTTTGGGTAAATCGAG	GCGCTCTCGCATCTATTGTCTGTTC	CAAGGTTGCTCAAA 2847
Qy	2997	GTCTAGCTGTTGATACA	AGCTTCAACCGTGACATCAAAAGGTTGAGATGATGAGAG	3056
Db	2848	GTTCTAGCCGGAGTTA	CACAGAAATTTCAAGTCACTTCAAAAGCAGACGATGAGCT	2907
Qy	3057	TTCTCAGAGCTATATCA	TCTCAAATGGA	CTACCTTATTTGATACTCTACCACTTCTT 3116
Db	2908	TTCTCTGAGCTTTACAT	CTTCAAGTGA	CAACTTGTGATTCTCCGACAACACTTCTG 2967
Qy	3117	CTATTGAACCTTCATT	GTGTGTGCTGCTGGCGTTTCAATGGGATCAATAACGGATATGAG	3176
Db	2968	ATCATTAACATCATTT	GAGTTATGTGCGCGTTTCTGATGCCATTAGCAATGGCTATGAC	3027
Qy	3177	TCAATGGGGCCCCCT	CTTTGGGAGCTATTCTTTGCATTTGGGTGATGTGCCATCTTAT	3236
Db	3028	TCAATGGGGA	CTCTCTTTGGGAGACTTTTCTCGCTCTTTGGGTCA	TGTTCATTATAC 3087
Qy	3237	CCCTTTCTCAAAAG	TTGGTTGGAAGCCAAAACAGACACCAACGATTTGTCA	TCGCTGG 3296
Db	3088	CCATTCCTCAAGG	AATGCTTTGGGAGCAAGACAAAATGCTTACGATTAT	TGTGTGTGG 3147
Qy	3297	TCCATTCCTGCTG	CTTCAATCTTCTGCTCTTTGGGTTCCGATTGATCCTTCTTGGG	3356

Db 3148 TCTATTCTCTAGCTTCGATCTTGACACTCTTGTGGGTACAGATTAAACCCGTTTGCGCT 3207

Qy 3357 AAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTGGATTG 3397

Db 3208 A---AAGGGGACCACTGTTGGAGATCTGTGGTCTGAATTG 3245

RESULT	9
AR267558	
LOCUS	AR267558
DEFINITION	Sequence 7 from patent US 6495740.
ACCESSION	AR267558
VERSION	AR267558.1 GI:29697634
	3828 bp DNA linear PAT 10-APR-2003

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
1	(bases 1 to 3828)			
	Arriola, A., Williamson, R.E., Betzner, A.S. and Peng, L.	Manipulation of cellulose and/or .beta.-1,4-Glucan	Patent: US 6495740-A 7 17-DEC-2002;	Location/Qualifiers
				Unknown.
				ORGANISM

source	1. .3828	/organism="unknown"
BASE COUNT	1056 a	721 c 912 g 1139 t
ORIGIN		

Query Match	36.3%	Score 1429.4;	DB 6;	Length 3828;
Best Local Similarity	65.8%;	Pred. No. 4.8e-266;		
Matches 2172; Conservative	0;	Mismatches 1066;	Indels 63;	Gaps 4;

QY	124 GAAGTCGAGGGGAGGAAGCGATGCAGCGAGCGCCGGCTGTGGCCGGCTCCACAACC	183
Db	219 GATTCCGTGTAGAAGACATCATGAACTACTGTGTGTGGCTCATTGCTGGCTTCAACAACA	278
QY	184 GCACGAGCTCGTCATCCGCCGCCGACGCCGATCCCGGGCCGAAGCCGCCGGGAGC	243
Db	279 GAAACGAA TTCGTTCTCATTAACGCCGATGAGAGTGCCAGATAAGATCAGTACAGAAAC	338

[illegible]

**QY** 364 GCCGGAGGGCACGCAGAAGCTCCCCAGTGCAAGACTCGATACAAGCGCTTCAAGGCT 423  
| | | | | | | | | | | | | | | | | | | | |  
**Db** 459 GTAGAGAAGAAATCAAGCTGTCTCTCAGTCAAAACTCGATACAAAAGATTAAAGTA 518  
| | | | | | | | | | | | | | | | | | | | |  
**QY** 424 GCCAAGTTGACCCGTTGACGAGAGAGAACGCCGTCGATGACCTTGACAACGAATTCA 483  
| | | | | | | | | | | | | | | | | | | | |  
**Db** 519 GTCACGGCGTTGATGGAGATGATGAAGAGAAGACATTTGATGATCTTGAAGTAGT 578  
| | | | | | | | | | | | | | | | | | | | |

OY 484 ACTGGACGCGCCATGACTGCAGTCTGTGCCGAGTCATGCTTTACGGCCACATGAGCT 543  
||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 579 TTGATCATGGGATGAGACCCTGAACATGCGCGTGAAAGCGCACTCTTTCAGCGCCTTAACA 638  
  
OY 544 ACGGCGGTGGAGGTGACCCCTAATGGCGCGCCACAGAAGTTCCAGCTCAACCCCAATGTC 603  
||||| ||| ||| ||| ||| ||| ||| |||  
Db 639 CCGGTGCTGGTGGATTGGATTGAGCTCACCTCGGCTCT-----CAGATTC 683

Qy      604 CACTCTCAACCAACGGGAAATGCTGATGACATCCACCAGGAGCAGCAGCGCTGTGC    663  
       ||| || | | | | | | | | | | | | | |  
Db      684 CTCCTTGAATTATTTGATGAGAATGCTGATATGTAATTCGATCGTCATGCTCTTATCG    743  
  
Qy      664 CTTCCTCATGGGTGTGGGGAAGAAGATACATCCCCTCTTATGCCGATCCAGCT    723  
       || | | | | | | | | | | | | | |  
Db      744 TGCCTCTTCAACGGGATATGGGAATCGCGTCTATCTCGAACCGTTACAGATTCTTCTG    803

724 TACCTGTGCAACCCAGGCTATGTGACCACATCCAAAGATCTTGCTGCATATGGGTATGCTA 783

Db 804 CACCTCCACAGCGGAGATCAATGGTTCCTCAGAAAGATATTCGGAAATATGGTTATGGAA 863  
QY 784 GTGTGCTTGGAGGAACGGATGGAGAATTGGAGCAGAGACAAGAGAGATGCCAGCA 843  
Db 864 GTGTGCTTGGAGGACCCTATGGAAAGTTTGGAGAGACGACAAGCGGAAAAGCTTCAAG 923  
QY 844 CGGGGAATGATGTGTGGTGGATGATGT-----GACG 876  
Db 924 TCATTAAGCATGAAGAGGAAACAATGTGCGAGTTCCAAATGATGACGACGAACTAGATG 983  
QY 877 ATGCTGATCTACCACCTAAATGGATGAAGCAAGACAACACTGTCCAGAAAATTCACCTTC 936  
Db 984 ATCCTGACATGCCCTATGATGATGAAGAGACACAACCTCTCTCAAGAAAGCTACCTATTC 1043  
QY 937 CATCAAGCCAGATTAAATCCATATAGGATGATTATCATTTATTCGGCTTGTGCTTTGGGT 996  
Db 1044 GTTCAAGCAGAAATAATCCTTACAGGATGTTAATTCGTGTCGCCCTCGCGATTCTTGGTC 1103  
QY 997 TCTTCTTCCACTACCGAGTGAATGCATCCGGTGAATGATGCATTTGCTTGTGGCTCATAT 1056  
Db 1104 TTTTCTTCAATTATAGAAATCTCCATCCAGTCAATGATGATATGATTATGGTTAAAGT 1163  
QY 1057 CTGTATCTGTGAATCTGTGTTGCCATGTCTTGATCTTGATCAATTTCCAAAGTGT 1116  
Db 1164 CAGTTATATGCGAAATATGTTGACAGTGTCTTGATCTTGATCAATTTCCCAAATGTT 1223  
QY 1117 TCCCTATTGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTGACAAAGAGGCC 1176  
Db 1224 ATCCTATAGAACGTGAACATACCTCGATAGACTCTCTCAGGTACGAGAGAGAAAGAA 1283  
QY 1177 AGCCATCTCACTTGTCCCAATTGATTTCTTGTCAATACGGTTGATCCCTTAAAGAAC 1236  
Db 1284 AACCCTCAGGATTAGCACCTGTGATGTTTGTGTAGTACAGTGGATCCGTTGAAGAAGC 1343  
QY 1237 CTCCCTTGTGCACAACAATACTGTTCTATCTATCTCTTTCGGTGATATCTCTGTGATA 1296  
Db 1344 CCCCCTGATTACAGCAACAACAGTCTTCCATTTAGCAGTGAATTAATCTCTGTGATA 1403  
QY 1297 AGGTTTCTGCTATGTTTCTGATGATGGTGTGCAATGCTAAAGTTGAAGCATATCTG 1356  
Db 1404 AGGTTGCTGTTATGATTAACAATGCTGAGCTATGCTTAATTTGAAGCTCTCTCTG 1463  
QY 1357 AAACATCTGAATTTGCAAGAATAAGGTTCTTTCGCAACGCTACATATGAACCTC 1416  
Db 1464 ATACAGCTGATTTGTCTACAAATAGGTTCTTTTGTAAAGATTTAATATCGAGCCAC 1523  
QY 1417 GCGCTCCAGAGTGTACTTCCACAGAGATAGACTACTTGAAGAACAAGGTGGCAGCAA 1476  
Db 1524 GAGCTCCTGAGTGTATTTTCTCAGAAGATGATTAACCTGAAGAACAAGTTCACTCTG 1583  
QY 1477 ACTTTGTAGGAGAGAGCAATGAAGAGAGATGAAGAAATTCAGAGTGAATCA 1536  
Db 1584 CTTTGTCAAGGAAACGTCGTCTATGAAGAGAGATTGAAGAGTTTAAAGTGAAGATTA 1643  
QY 1537 ATGCTTATGTTGCCAAAGCCAGAAAGTTCTCTGAAGAGAGATGACAAATGCAAGATGGAA 1596  
Db 1644 ATGCACTGTTGCTACTGCAAGAAAGTCCCTGAGGAAAGTTGACTATGCAAGATGGAA 1703  
QY 1597 CCCCCTGGCCTGAAAACAATGTTGCTGATCATCTCTGAATGATTCAGGTCTTCTTGGCC 1656  
Db 1704 CTCCCTGGCCTGAAAACAACGTCGTGACCATCTGGAATGATTCAAGTGTCTTGGGTC 1763  
QY 1657 AAAGCGAGGCTTGACTGAGGGAATGAACCTGCCAGATTGGTTATGTTCTAGAG 1716  
Db 1764 ATAGTGAGATTGATACGGAATGTAATGAGTTACCAAGTCTAGTATGTTCTCGTG 1823  
QY 1717 AGAAACGACGAGCTATAACCATCATAGAAGAGCTGTGCTATGAATGCATTGTGTCGAG 1776  
Db 1824 AGAAGCGGCTGATTTGATCAACCAAGAAAGCTGAGCTATGAATTCCTTGATTCGAG 1883  
QY 1777 TCTCTGCTGTACTACAAATGCTCCATATTTGTTAACTTTGATTTGATCACTACATCA 1836  
|||||

Db 1884 TCTCTGCTGTTCATCAAAAGCTCCTTACTTCTTAATGTGATTTGATCACTACATCA 1943  
QY 1837 ACAACAGCAAGGCTATAAAGAGCAATGTGTTTATGATGACCCCTTACTAGAAAGA 1896  
Db 1944 ACAACAGCAAGCAATTAGAGATCTATGTCTTATGATGACCCGCAATCGGAAAGA 2003  
QY 1897 AGTTTGTCTATGACAGTTCCCTCAAGATTGATGGGATTTGATCGCATGACCGATATG 1956  
Db 2004 AAGTTGTATGTTCAGTTTCCGACAGATTGATGGGATTTGATGACATGATAGATACT 2063  
QY 1957 CTAAACCGAATGTGCTTTTGTGATATCAATGAAGGTTTGATGTTGATTCAGGCTC 2016  
Db 2064 CAAACCGTAACGTGTGTTCTTGATATTAACATGAAGGTTCTTGATGGATTAACAAGGAC 2123  
QY 2017 CAATTATGTTGCTACTGATGTGATTTAGAAAGCAGGCAATTAATGTTATGATGCC 2076  
Db 2124 CGATATATGTCGGGACAGGTTGTGTGTTAGAAAAAGGCTCTTATGTTTGTGATGCAC 2183  
QY 2077 CCAAAACAAGAACCCATCAAGACTTGCAACTGCTGCGCCCAAGTGTCTTTGCT 2136  
Db 2184 CAAAGAAAGAAACCAACGCAAAACCTGTAACGTGTGCGCTTAATGTGTGTTGT 2243  
QY 2137 GTTGTGCTTTGGCAATAGGAAGCAAAAGACATAACCAACCCAAACAGAGAAAGAAA 2196  
Db 2244 GTTGTGGTTGAGAAAGAGATAAACGAAACCCACAGATTAAGAAACTAACACTAAAG 2303  
QY 2197 AGTTATTTTTCAGAAAGAGAACCAATCCCTGCATATGCTCTGTGTAATTTG 2256  
Db 2304 AGACTTCA-----AAGCAGATTATGCGCTAGAGAAATGTGACGAAAG 2345  
QY 2257 ACGAAGCTGCTCCAGGAGCTGAGAAATGAAGCCGGTATTTGTAATCAACAAAATTAG 2316  
Db 2346 GTGTATGCTCCAGTGTCAAAATGTGAGAAAGATCTGAAGCAACAAATTGAATTTGG 2405  
QY 2317 AAAAGAAATTTGCCCAATCTTCTGTTTGTATCATCCACACTTCTGAGAAATGTGAA 2376  
Db 2406 AGAAGAAATTTGACAATCTCCGTTTGTGCTCTGCTGTTCTACAGAACGCTGAG 2465  
QY 2377 CCTGAAGAGTGAAGTCCCTGCTCTCTTTGAAGAAGCTATACATGTCATTAGTTGTG 2436  
Db 2466 TTCCCGGTAAACGCAAGCCCCGATGTTGTTAAGAGAAGCCATTCAAGTTATTAAGTGG 2525  
QY 2437 GTTATGAAGACAGACAGACTGGGGAAGAGATTGGCTGATCTATGATCAGTTACAG 2496  
Db 2526 GGTACCAAGATTAACCGAATGGGGAAGAGATCGGGTGAATTTATGATCGGTGACTG 2585  
QY 2497 AAGATATTCTAATGCTGTTCAAGATGATGTCATGTTGGCGGTCAAATTTACTGCATAC 2556  
Db 2586 AAGATATCTGACGGGTTTCAAGATGATGCTCATGATGAGATCTGTATCTGTATGC 2645  
QY 2557 CTAACCGGTTGCAATCAAAAGTTCTGCACTCTGAATCTTTCAGATCGTCTTCAACGAG 2616  
Db 2646 CTAAGCGTGACGTTTAAAGATCTGCTCTATTAACCTGTGAGATCGTCTTCAACAAG 2705  
QY 2617 TGCTTGGTGGCTCTTGGGCTATGAGATCTTCTTCAAGCAATCATGCCCCCTTTGGT 2676  
Db 2706 TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTCTTGAAGACAGATGTCCGATATGCT 2765  
QY 2677 ATGGTATGTTGGCGCTCGAAATTTTGAAGAAAGATTTTCTCATCATCACTCATCTGCT 2736  
Db 2766 ATGTTATGTTGGTGTTTAAATGCTTGAAGAGATTTCTTACATCAACTCTGTCTCT 2825  
QY 2737 ATCTTGACATCTATTCCTCTTGGCTTACTGTACATTTGCTGCACTGTCTTATTTGA 2796  
Db 2826 ATCTTGACATCTTCACTTCAATGATGCTATGTTCTCTCCCGGTTGTGTTACTCA 2885  
QY 2797 CAGGAATTTATCACTCCAGAGCTGAATTAATGTTCCAGCCGTGTGTTCACTGCTT 2856  
Db 2886 CAGGAATTTATCATGTCCTGAGATTAAGCACTACGAGGTATACTCTTCACTGCTCATGT 2945  
QY 2857 TTATCTGATTTTCTCTACGAGCATCTGAAATGAAGATGAGTGAAGTGTGTAATTTGATG 2916  
Db 2946 TCATATCCATGACGATTAAGTGAATCTGGAATCTGGAATGCAATGGGAGGTGCGGAATCATG 3005  
|||||



[illegible]

```

RESULT 10
AX030944
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
CDS

```

AX030944 3828 bp DNA linear PAT 24-NOV-2000

Sequence 7 from Patent WO9800549.

AX030944 AX030958

AX030944.1 GI:10278347

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

Williamson,R.E., Peng,L., Arioli,A. and Betzner,A.S.

Manipulation of cellulose and/or beta -1,4-glucan

Patent: WO 9800549-A 7 08-JAN-1998;

WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ; COMMW SCIENT IND RES ORG (AU)

On Oct 15, 2002 this sequence version replaced gi:10278361.

Location/Qualifiers

1..3828

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/strain="COLUMBIA"

/db\_xref="taxon:3702"

/clone="ATH-A"

239..3493

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC09626.1"

/db\_xref="GI:10278348"

/db\_xref="REMTREMBL:CAC09626"

/translation="MNTGRLLAGSHNRNEFLVINADESARIRSVQELSGTQICGD  
EIELTVSSELFVACNECAFPPVCRPCYEYERREGNACPOCKTRYKRIKGSPPVDGDDE

EEBEDIDLEVEFDHGMDEPHAAEALSSRINTGRGGLDSAPPGSQIPLTTYCDEBADM  
 YSDRHALIVPSTGYGNRVYPAFTDSSAPQARSMVPQKDIAYGYGSVAMKDRMEV  
 MKRQGEKLOVYI KHEGNNNGRSDNDDLELDPDMPEMDEGRQPLSRKPIRSSRINPY  
 RMLILCRALAIGLFPHYRILHPVANDAYGLMLTSVICEIMFAWSWILDOPEKMPIREB  
 TYLDRLSLRYEKEGKPSGLAPVDVFSTVDPLKEPPLITANTVLSILANDY PVDKVAC  
 YVSNNGAAMLTEALSDTADFATKWVPCKKFNIEBPRAPEWFSQKMDYLKXVHPAF  
 VRERAMKRDYEEFKVINATALVATAQKVEERWTMODGTPMGNNVRDHGMIQVFLG  
 HSGVRDITDGNELPRLVYVSREKRGFDHKKAGAMNSLIRVASVNLNAPYLLNVDCDH  
 YINNSKAIRESMCFEMMDPQSGKKVCYVQFQRFDDIRDRYSNRNVFPDINMKGLD  
 GIQGPYVGTGCVFRKQALYGFDA PKKKPBGKTCNCPWKCCLCGRKSKXTKATD  
 KKTNTKETSKQIHAIENDEGVI VPSVNEKRSEATQULKLEKKFGQSPVVASAVLQNN  
 GGVPRNASPACILREAOVISCSGYQDKTEKGIEIMIGYSTEDILTFGRWCHGRSS  
 VYCMFKRAAFKGSAPINLSDRHLQVLRNALGSVEIFLSRHCPIMWYGGGKMLERFS  
 YINSVVYPMTSLPIVYCSLPVAVCLTGKFI VPEISNAGILFMLMFISIAVTGILEM  
 QMGVGVIDDWMRNEQFWIVIGASSHLFALFOGLIKVLAVGNTNFTVTSKADDDGAFSEE  
 LYIFKMTTLIPPTLLIINIGIVGVSADAINSGYDSWGPFLGRLFALMWIVHLXP  
 FLKGM LGQDKMPTIIVVMSILASILTLTLWVRINPFVAKGGPVLEICGLNCGN"

[illegible]

Db	924	TCATTAAGCATGAAGAGAAAACAATGTCGAGTTCCAAATGATGACGACGAACTAGATG	983
Qy	877	ATGCTGATCTACCCTAATGATGAAGCAACAACACTGTCAGGAAATTCACCTTC	936
Db	984	ATCCTGACATGCCCTATGATGATGAAGGAAGACAACTCTCTCAAGAAAGCTTACTTTC	1043
Qy	937	CATCAAGCCAGATTAACTCATATAGGATGATTATCATTTTCGGCTGTGGCTTTGGGGT	996
Db	1044	GTTCAAGCAGAAATAATCTTACAGAGATGTTAATTTCTGTGTCGCTCGCATTTGGTTC	1103
Qy	997	TCTTCTTCACCTACCGAGTATGCATCCGGTGAATGATGATTTGCTTTGGCTCATAT	1056
Db	1104	TTTTCTTCATTATAGAAATTCATCCATCCAGTCAATGATGATGATTATGGTTAACGT	1163
Qy	1057	CTGTTATCTGTGAATCTGCTTGGCCATGTCCTTGATTTCTGATCAATTTCCCAAGTGT	1116
Db	1164	CAGTTATATGCGAAATATGGTTGACAGTGTCTTGATTTCTGATCAATTTCCCAATGGT	1223
Qy	1117	TCCCTATTGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGTTTCGACAAGAGGCC	1176
Db	1224	ATCCTATAGAACGTGAACATACCTCGATAGACTCTCTCAAGTACGAGAAGAGAA	1283
Qy	1177	AGCCATCTCAACTTGCTCCAAATTGATTTCTTGTGACATCGGTTGATCCCTTAAAGAAC	1236
Db	1284	AACCGTCAGATTAGCACCTGTGATGTTTGTGTAGTACAGTGATCCGTTGAAAGAGC	1343
Qy	1237	CTCCTTTGGTCACAAACAAATCTGTTCTATCTATCTTGGTGATTAATCCTGTTGATA	1296
Db	1344	CCCCCTTGATTACAGCAACACAGTCTTCCATTTACAGAGTTGATTATCCGTGGATA	1403
Qy	1297	AGTTTCTTGCTATGTTTCTGATGATGGTGTGCTCAATGCTAACGTTGAAGCATTAATCTG	1356
Db	1404	AGGTGGCGTGTATGATCAACAATGGTGACGTATGCTTACATTTGAAGCTCTCTCTG	1463
Qy	1357	AAACATCTGAATTTGCCAABGAATGGGTTCCCTTTCGAAACGTTACAATATTGAACCTC	1416
Db	1464	ATACAGCTGATTTTGCTACAAATAGGGTTCCTTTGTAGAAGTTTAATATCGAGCCAC	1523
Qy	1417	GCGCTCCAGAGTGTACTTCCACAGAGATAGACTTCTGAAAAGCAAGGTGGCACA	1476
Db	1524	GAGCTCCTGAGTGTAATTTTCTCAGAAAGATGATTACTGGAAGAACCAAGTTCACTCTG	1583
Qy	1477	ACTTGTAGGAGAGAGAGCAATGAAGAGAGATGAGGAATTCGAAGTGAGATCA	1536
Db	1584	CTTTGTGAGGAACGTCGTCTATGAAGAGAGATTATGAAGATTAAGTGAAGATTA	1643
Qy	1537	ATGCTTTAGTTGCCAAAGCCGAGAAAGTTCCTGAAGAGATGGAACAATGCAAGATGAA	1596
Db	1644	ATGCACTGGTGTCTACTGCACAGAAAGTGCCTGAGGAACGTTGACTATGCAAGATGAA	1703
Qy	1597	CCCCCTGGCCTGGAACAATGTCGTGATCATCCTGGAATGATTCAGGCTTCTCTGGCC	1656
Db	1704	CTCCTTGGCCTGGAACAACGTCGTAACCATCCTGGAATGATTCAGGTGTTCTTGGGTC	1763
Qy	1657	AAAGCGGAGGCTTGACTGTGAGGGAATGAACCTGCCACGATTGGTTATGTTTCTAGG	1716
Db	1764	ATAGTGAAGTTCGTGATACGGAATGTAATGATTACCAACGTTAGTGTATGTTCTCGTG	1823
Qy	1717	AGAAACGACCAGGCTATTAACCATCATAGAAAGCTGGTGCTATGAATGCATTGGTCCGAG	1776
Db	1824	AGAAAGGCGCTGATTGATTCACCAAGAAAGCTGAGCTATGAATTCCTTGATCCGAG	1883
Qy	1777	TCTGTCTGTACTAACAATGCTCCATATTTGTTAACTTGATTTGTGATCACTACATCA	1836
Db	1884	TCTGTCTGTCTATCAACAAGCTCTTACCTTCTTAATGTGATTTGTGATCACTACATCA	1943
Qy	1837	ACAACAGCAAGGCTATTAAGGAAGCAATGTGTTTATGATGACCCCTTACTAGAAAGA	1896
Db	1944	ACAACAGCAAGCAATTAAGAAATCTATGTGTTTATGATGACCCGCAATCGGAAAGA	2003
Qy	1897	AGGTTTGTATGTACAGTTCCTCAAGATTGATGGGATGATCGCCATGACCGATATG	1956

Db	2004	AAGTTTGTTATGTTTCAGTTTCCGCAGAGATTTCGATGGGATTGATAGACATGATAGATACT	2063
Qy	1957	CTAACCCGAATGTGTCCTTTTGTGATATCAATGAAAAGTTTGGATGTTACAGGTC	2016
Db	2064	CAAAACCGTAACTGTGTCTTTGATATTAACTGAAAAGCTTGTATGGGATACAAAGAC	2123
Qy	2017	CAATTTATGTTGTACTGATGTGATTTAGAAAGCAGGCAATTAATGTTATGATGCC	2076
Db	2124	CGATATATGTCGGGACAGGTTGTGTGTTAGAAAAACAGGCTCTTATGTTTGTATGCAC	2183
Qy	2077	CCAAAAACAAGAACCCATCAAGACTTCAACTGCTGGCCCAAGTGTCTTTTGCT	2136
Db	2184	CAAGAAGAGAAGAACCAACGCAAAACCTGTAACGTGTGGCTAAATGGTGTGTTGT	2243
Qy	2137	GTTGCTGCTTTGGCAATAGGAAGCAAAAGAACTACCAAAACCCAAAACAGAGAAGAAA	2196
Db	2244	GTTGTGGGTTGCAAAAGAGATMAAACGAAACCCACAGATTAAGAAACTAACACTAAAG	2303
Qy	2197	AGTTATTATTTTTCAAGAAAGAGAACCAATCCCCCTGCATATGCTCTTGGTGAATTG	2256
Db	2304	AGACTTCA-----AAGCAGATTTCATGCGCTAGAGAAATGTCGACGAAG	2345
Qy	2257	ACGAAGCTGCTCAGAGACTGAGAAATGAAAAGCCGGTATTGTAATCAACAAAATTAG	2316
Db	2346	GTGTTATCGTCCCACTGTCAAAATGTTGAGAAGAGATCTGAAGCAACACAAATTGGA	2405
Qy	2317	AAAAGAATTTGGCCAATCTTCTGTTTGTATCATCCACACTTCTCGAAGTGTGGA	2376
Db	2406	AGAAGAAGTTTGGACAATCTCCGGTTTTCGTTCCCTCTGCTGTCTACAGAACGTTGAG	2465
Qy	2377	CCTGAAGAGTGCAGTCTCTGCTTCTCTTTGAAGAAGCTATACATGTCATTAAGTTGTG	2436
Db	2466	TTCCCCGTAACGCAAGCCCCCGCATGTTGTGTTAAGAGAAGCCATTCAAGTTATTAGCTCG	2525
Qy	2437	GTTATGAAGACACAGACAGACTGGGGAAGAGATTGGCTGATCTATGATCAGTTACAG	2496
Db	2526	GGTACCAAGATMAAACCGAATGGGGAAGAAAGATCGGGTGATTATGATCGGTGACTG	2585
Qy	2497	AAGATATTCTAACTGCTTCAAGATGCAATGTCATGGTTGGCGGTCAATTTACTGCATAC	2556
Db	2586	AAGATATCCTGACGGGTTTCAAGATGCATTCGATGAGATGAGATCTGTGTACTGTATGC	2645
Qy	2557	CTAAACGGGTGCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAACGAG	2616
Db	2646	CTAAGCGTGACGCTTTAAAGGATCTGCTCTTAATTAACCTGTGATCGTCTTCAATCAAG	2705
Qy	2617	TGCTTCGGTGGCTCTTGGGCTATTTGAGATCTTCTTCAGCAATCATTGCCTCTTGGT	2676
Db	2706	TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTCTTGAGCAGACATTTGTCGATATGCT	2765
Qy	2677	ATGGGTATGTTGGCGGCTCTGAAATTTTGGAAAGATTTTCTACATCAACTCCATCGTGT	2736
Db	2766	ATGTTATGTGTGTGTTTAAATGTTGAGAGATTCTCTTACATCAACTCTGTCTCT	2825
Qy	2737	ATCCTTGACATCTATTCCTCCCTCTTGCTTACTGTACATTTGCCCTGCATCTGTTATTGA	2796
Db	2826	ATCCTTGACTTCACTTCCATTTGATCGTCTATTTGTTCTCTCCCGGTTGTTTACTCA	2885
Qy	2797	CAGGAAATTTATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGTTCAATGCACTTT	2856
Db	2886	CAGGAAATTCATCGTCCCTGAGATTAAGCAACTACGCAAGTATACTCTTCATGCTCATGT	2945
Qy	2857	TTATCTGCATTTTGTCTACGAGCATCTAGAAATGAGATGGAAGTGTGTTGAATTGATG	2916
Db	2946	TCATATCCATAGCAGTACTGGAATCCTCGAAATGCAATGGGAGAGTGTGGAATCGATG	3005
Qy	2917	ACTGTTGAGGAATGAGCAGTTCTGGGTCAATTTGAGGTGTCTCTCACACCTCTTGTCTG	2976
Db	3006	ATTGTTGAGGAAACGACAGTTTTGGGTAATCGGAGGGGCTCTCCGATCTATTTGCTC	3065
Qy	2977	TGTTCCAGGACTTCTCAAGCTCATAGCTGTGTGTTGATACAAAGCTTCAACCGTGAATCAA	3036
Db	3066	TGTTTCAAGGTTTGTCTCAAGTCTAGCCGAGTTAAACGAATTTACAGTCACTTCAA	3125

QY	3037	AGCGTGAGATGATGAGGAGCTTCTCAGAGCTATATACATTCAAAATGAGACTACCTTATTGA	3096
Db	3126	AAGCAGCAGACGATGAGAGCTTCTCTGAGCTTTACATCTTCAAGTGAGACAACCTTGTGA	3185
QY	3097	TACCTCCTAACCACTTGCTCTCTATTTGAACCTTCATTGGTGTGCTGGCGTTTCAAAATG	3156
Db	3186	TTCTCCGACAACACTTCTGATCATTTACATCATTTGAGATTATTGTCGGCGTTTCTGATG	3245
QY	3157	CGATCAATTAACGATATAGATCATGGGGCCCCCTTTGGGAAGCTATTCTTGCATTTT	3216
Db	3246	CCATTAGCAATGGCTATGACTCATGGGAGCCCTCTTTGGGAGACTTTTCTTGCCTCTTT	3305
QY	3217	GGGTGATTGTCCATCTTATATCCCTTCTCAAAGTTTGGTTGGAAGGCAAAACAGACAC	3276
Db	3306	GGGTCAATTGTTCAATTATTAACCCATTCCCTCAAGGGAATGCTTGGGAAGCAAGACAATAATGC	3365
QY	3277	CAACGATTGTCAATCGTCTGGTCCATTCTGCTGGCTTCAATCTTCTCGCTCCTTTGGGTTTC	3336
Db	3366	CTACGATTATGTGTGTCTGTCTATTTCTTCTAGCTTCGATCTTGACACTCTTGTGGGTCA	3425
QY	3337	GGATTGATCCTTTCCTTGGCAAGGATGATGATGATCCGCTTCTTGAGGAGTGTGTTGGATT	3396
Db	3426	GAATTAACCCGTTTGTGGCTA--AAGGGGAGCCAGTGTTGGAGATCTGTGTCTGAATT	3482
QY	3397	G 3397	
		—	
Db	3483	G 3483	

RESULT	11		
BD022677			
LOCUS	BD022677	3828 bp	DNA linear
DEFINITION	Manipulation of cellulose and/or beta-1,4-glucan.		
ACCESSION	BD022677		
VERSION	BD022677.1	GI:22563900	
KEYWORDS	JP 2001510326-A/5.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 3828)		
REFERENCE	Ariori, A., Williamson, R.E., Bettsunar, A.S. and Penn, L. Manipulation of cellulose and/or beta-1,4-glucan Patent: JP 2001510326-A 5 31-JUL-2001;		
AUTHORS			
TITLE			
JOURNAL			

COMMENT	OS	PN
	Arabidopsis thaliana (thale cress)	JP 2001510326-A/5

PD 31-JUL-2001  
PF 24-JUN-1997 JP 1998503665  
PR 27-JUN-1996 AU PO 0699  
PI ANTONIO ARIORI, RICHARD EDWARD WILLIAMSON, ANDREAS STEPHAN PI  
BETSUNAR,  
PI LYONKAI PENN  
PC C12N15/09, A01H5/00, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/  
PC 10, C12N5/10,  
PC C12N9/10, C12N15/00, C12N5/00, C12N5/00  
CC Strandedness: Single;  
CC Topology: Linear;  
FH Key location/Qualifiers  
FT CDS 239..3490.  
Location/Qualifiers

FEATURES	source	Location/Qualifiers
	1.	3828
		/organism="Arabidopsis thaliana"
		/mol_type="genomic DNA"
		/db_xref="taxon:3702"
BASE COUNT	1056 a	721 c 912 g 1139 t
ORIGIN		

Query Match	36.3%;	Score 1429.4;	DB 6;	length 3838;
Best Local Similarity	65.8%;	Pred. No. 4.8e-266;		

Matches 2172; Conservative 0; Mismatches 1066; Indels 63; Gaps 4;

QY	124	GAA	TGAGGGGAGGAGAACGATGGAGGGCGAGCGCCGGGCTGTGGCCGGCTCCCAACAACC	183
Db	219	GATT	CGGTAGAGACATCATGAATACTGGTGGTCGGCTCATTTGCTGGCTCTCACACA	278
QY	184	GCA	ACGAGCTCGTCGTATCCGCCCGACGGCGATCCCGGCCGAAGCCGCCGGAGC	243
Db	279	GAA	CGAATTGTTCTCATTAACCGCGATGAGAGTCCAGATAAGATCAGTACAGAAGC	338
QY	244	AGA	ACGGGCGAGTGTGCGAGATTTGCGGCGACGACGTGGCTTGCCCCGGCGGAGC	303
Db	339	TGAG	TGGGCAACATGTCAATCTGTGAGATGAATCGAATTAACGGTTAGCAGTGAGC	398
QY	304	CCT	CGTGGCGTGCACGAGTGGCCCTTCCCGCTGCGCGGACTGCTACGAATACGAGC	363
Db	399	TCT	TGTGCTTGCCACGCAATGCCCATTTCCCGGTTGTAGACCATGCTATGAGTATGAAC	458
QY	364	GCC	GGGAGGGCAGCAGAACTGCCCCAGTGCAGACTCGATACAAGCGCCTCAAGGCT	423
Db	459	GTA	GAGAGGAATCAAGCTTGCTCAGTGCAAAACCTCGATACAAAAGATTAAAGTA	518
QY	424	GCCA	CGTGTGACCGGTGACGAGAGGAGGACGCGGTGATGACTTGACAACGAGTTCA	483
Db	519	GTCC	ACGGTGTATGGAGATGATGAAGAAGAAGACATTGATGATCTTGAGTATGAGT	578
QY	484	ACT	GGAGCGCCATGATTCGAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT	543
Db	579	TGAT	CATGGATGAGACCTTGACATGCGCTGAAGCCGCACTCTTTCACGCCCTTAACA	638
QY	544	ACG	CGCGTGAGGTGACCTTAATGGCGCGCCACAAGCTTTCAGCTCAACCCCAATGTT	603
Db	639	CCG	GTCGTGGTGGATTGATTCACTCCACCTGGCTCT-----CAGATTTC	683
QY	604	CAC	TCTCACCACCAACGGGCAATGCTGATGACATCCACCGAGCAGACGCGCTGTGC	663
Db	684	CTCT	TGTACTTATTTGATGAAGATGCTGATATGATTTGATCGTCATGCTCTTATCG	743
QY	664	CTT	CTTTCATGGGTGGGGGAAAGAGATACATCCCTTCCTTATGCGGATCCAGCT	723
Db	744	TGC	CTCCTTCAACGGGATATGGGAATCGGCTATCTCTGACCGCTTACAGATTCTTCTG	803
QY	724	TAC	CTGTGCAACCCAGTCTATGGAACCCCATCCAGAGATTTGCTGCATATGCGTATGCTA	783
Db	804	CAC	CTCCACAGGCGAGATCAATGCTTCTCAGAAAGATATTCGGAATATGTTATGGAA	863
QY	784	GTT	GCTTGGTGGAGGACGATGGAGAATTGGAGCAGAGCAAGAGAGATGACCAAGAG	843
Db	864	GTT	GCTTGGAGGACCGTATGGAGTTTGGAGAGACGACAAAGCGAAAGCTTCAAG	923
QY	844	CGG	GGAATGATGTTGTTGATGATGT-----GAGC	876
Db	924	TCAT	TAAAGCATGAAGAGGAAACAATGTCGAGGTTCCAATGATGACGACGAAGTAGATG	983
QY	877	ATG	TGATCTACCACTAATGATGGAAGCAAGACAACAACACTGTCCAGAAAAATTCACCTTC	936
Db	984	ATC	TGACATGCTTATGATGATGAAGGAAGACAACCTCTCTCAAGAAAGCTACTATTTC	1043
QY	937	CAT	CAAGCCAGATTAACTCATATAGATGATTATCATTAATTCGGCTGTGCTTTGGGGT	996
Db	1044	GTT	CAAGCAGATAAATCTTACAGAGTGTAAATTTGTGTGCTCGCGATTCTTGGTC	1103
QY	997	TCT	TCTTCCACTACCGAGTATGATCCGCGTGAATGATGCAATTGCTTTGTGCTCATAT	1056
Db	1104	TTT	CTTTCATTATAGAAATCTCCATCCAGTCAATGATGCAATATGATTATGTTAACGT	1163
QY	1057	CTG	TATCTGTAATCTGGTTTGGCATGTCTTGATTTTGATCAATTCCCAAGTGGT	1116
Db	1164	CAG	TATATGCGAAATATGTTTGCAGTGTCTTGGAATTTTGATCAATTCCCAATATGT	1223
QY	1117	TCC	CTATTGAGAGAGACTTAACCTAGAACCGGCTGTCACTGAGGTTGCACAAGGAAGCC	1176
Db	1224	ATC	CTATAGAACGTGAACATACCTCGATAGACTCTCTTCAGAGTACGAGAGGAAGGAA	1283



QY	1177	AGCCATCTCAACTGCTCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGGAAC	1236
Db	1284	AACCGTCAGGATTAGCACTGTGTGATGTTTTTTGTTAGTACAGTGGATCCGTTGAAAGAGC	1343
QY	1237	CTCCTTTGGTTCACAACAATACTGTTCTTATCTATCCCTTCCGGTGATTAATCCTGTGATA	1296
Db	1344	CCCCCTGATTACAGCAAAACACAGTCTCTTCCATTCTACAGTTGATTAATCCTGTGATA	1403
QY	1297	AGGTTTCTTGCTAATGTTTTCTGATGATGTCGTCGAATGCTAAGCTTTGAAGCATTAATCTG	1356
Db	1404	AGGTGCGGTATATGATATCAACAATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG	1463
QY	1357	AAACATCTGAATTTGCAAAAGAAATGGGTTCCTTTCTGCCAAACGGTACAAATATTGAACCTC	1416
Db	1464	ATACAGCTGATTTTGTCTACAAAATGGGTTCCITTTTGTAAAGAAGTTTAATATCGAAGCCAC	1523
QY	1417	GCGCTCCAGAGTGTACTTCCAAACAGAGATAGACTACTTGAAGAACAAGTGGCAGCAA	1476
Db	1524	GAGCTCCTGAGTGTATTTTCTCAGAAAGATGATTAACCTGAAGAACAAAGTTTCATCCTG	1583
QY	1477	ACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATGAGGAATTCAGGATTCAGAGATCA	1536
Db	1584	CTTTTGTACGGAGACGTCGTCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA	1643
QY	1537	ATGCCTTAGTTGCCAAAGCCAGAAAGTTCCTGAAAGAGATGGACAATGCAGATGGAA	1596
Db	1644	ATGCACTGTTGCTACTGCAAGAAAGTCCCTGAGGAAAGTTGGACTATGCAAGATGCAA	1703
QY	1597	CCCCCTGGCCTGAAACAATGTCGTGATCAATCCTGAATGATTCAAGTCTTCTCTGGCC	1656
Db	1704	CTCCTTGCCCTGAAACAACGTCCTGACCATCCTGGAATGATTCAAGTGTCTGGGTC	1763
QY	1657	AAAGCGAGGCGCTTGACTGTGAGGAAATGAACGTCGCCAGATTGTTATGTTTCTAGAG	1716
Db	1764	ATAGTGAGAGTTCGTGATACGGATGTAATGAGTTACCAAGTCTAGTATGTTCTCGTG	1823
QY	1717	AGAAACGACCAGGCTATAACCATCATTAAGAAAGCTGTGCTATGAATGCATTGTCGGAG	1776
Db	1824	AGAAAGCGGCTGGAATTGATCAACCAAGAAAGCTGGACATGAATTCCTTGATCCGAG	1883
QY	1777	TCTCTGCTGTACTAACAAATGCTCCATATTTGTTAACTTGGATTGTGATCACTACATCA	1836
Db	1884	TCTCTGCTGTCTATCAACGCTCCTTACCCTTTAATGTGCAATGTGATCACTACATCA	1943
QY	1837	ACAAACGACAAGCCTATTAAGSAAAGCAATGTGTTTATGATGGAACCCCTTACTAGSAAAGA	1896
Db	1944	ACAACAGCAAGCAATTAGAAATCTATGTGTTTCATGATGGAACCCGCAATCGGGAAGA	2003
QY	1897	AGGTTTGCTATGTACAGTTCCCTCAAGATTTGATGGAAATGATCGCCATGACCGAATATG	1956
Db	2004	AAGTTTGTATGTTCAGTTTCCGACAGAGATTGTATGGGAATTGATAGACATGATGATACT	2063
QY	1957	CTAACCCGAATGTGTCTTTTGTATATCAACATGAAAGGTTTGGATGTGATTTCAAGGTC	2016
Db	2064	CAAAACCGTAACGTTGTGTTCTTTGATATTAACATGAAGGTCCTTGATGGATACAGSAC	2123
QY	2017	CAATTTATGTTGTTACTGATGTGTATTTAGAAAGCAGGCATTAATGTGTTATGATGCC	2076
Db	2124	CGATATATGTCGGACAGGTTGTGTGTTAGAAAACAGGCTCTTATGTTTGAATGAC	2183
QY	2077	CCAAAACAAGAGCCACCATCAAGGACTTGCAACTGTCGCCCCAAGTGTCCTTTTCT	2136
Db	2184	CAAGAAGAAAGAAACCAACGCAAAACCTGTACTGTGGCCTAATGTGTGTTTGT	2243
QY	2137	GTTGCTGCTTTGGCAATAGGAAGCAAAAGAGACTACCAAAACCAAAACAGAGAAAGAA	2196
Db	2244	GTTGTGGGTTGAGAAAGAGAGTAAACGAAAGCCAAGATTAAGAAACTAACACTTAAG	2303
QY	2197	AGTTATTAATTTTCAAGAAAGAGAAACCAATCCCTCGCATATGCTCTTGTTGAATTG	2256
Db	2304	AGACTTCA-----AAGCAGATTCAATGCGCTAGAGAAATGTCGACGGAAG	2345

Qy	2257	ACGAAGCTGCTCCAGAGAGCTGAGAAATGAAAGCCCGGTATTGTAAATCAACAAAATTAG	2316
Db	2346	GTGTTATCGTCCAGTAGTCAAAATGTTGACAGAGAGATCTGAAGCAACACAAATTGAAATTGG	2405
Qy	2317	AAAAGAAATTTGGCCAAATCTTCTGTTTTTTGTTACATCCACACTTCTCGAAGATGGTGAA	2376
Db	2406	AGAAGAAAGTTGGACAATCTCCGGTTTTCGTTGCCTCTGCTGTTCTACAGAACGGTGGAG	2465
Qy	2377	CCTTGAAGAGTCAAGTCCCTGCTTCTCTTTGAAAGAGCTATACATGTCATTAGTTGTG	2436
Db	2466	TTCCCCGTAACGCAAGCCCCCGCATGTTTGTTAAGAGAAAGCCATTCAAGTTATTAGCTGCG	2525
Qy	2437	GTTATGAAGACAGACAGACTGGGAAAGAGATTGGCTGGATCTATGGATCAGTTACAG	2496
Db	2526	GGTACCAAGATAAAACCGAATGGGAAAGAGATCCGGTGATTTATGATCCGGTGACTG	2585
Qy	2497	AAGATATTCTAACTGCGTTTCAAGATGCATTGTCTATGCTTGGCCGTCAAATTACTGCATAC	2556
Db	2586	AAGATATCCTGACGGGTTTCAAGATGCATTGCCATGGATGGAGATCTGTGTACTGTATGC	2645
Qy	2557	CTAAGCGGGTTCATTCAAAGGTTCTGCACCTCTGAATCTTCAAGATCGTCTTCAACGAG	2616
Db	2646	CTAAGCGTGCAGCTTTTAAAGGATCTGCTCTATTAACCTGTCAAGATCGTCTTCAACAG	2705
Qy	2617	TGCTTCGGTGGCTCTTGGGCTATTTGAGATCTTCTCAGCAATCATTTGCCCTTTGGT	2676
Db	2706	TTCTACGTTGGCTCTTGCGCTCTGTAGAGATTCTTGAAGACACATTTGCCGATATGGT	2765
Qy	2677	ATGGGATGCTGGCGGCTGAATTTTGGAAAGATTTTCCCTACATCACTCCATCGTGT	2736
Db	2766	ATGGTTATGGTGGTGGTTTAAATGTTGGAGAGATTCTCTTACATCACTCTGTCTCT	2825
Qy	2737	ATCCTTGACATCTATTCCTCCCTCTTGCGTTACTGTACATTTGCCCTGCCATCTGTTATTGA	2796
Db	2826	ATCCTTGACTTCACTTCCATTGATCGTCTATTGTTCTCTCCCGCGTTGTATTACTCA	2885
Qy	2797	CAGGGAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGTTCAATGTCATTT	2856
Db	2886	CAGGAATTCATCGTCCCTGAGATTAAGCAACTACGAGTATACTCTTCATGCTCATGT	2945
Qy	2857	TTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGAGTGTGTTGAATTGATG	2916
Db	2946	TCATATCCATAGCAGTAACCTGGAATCCTGAAATGCAATGGGAGGTGTCGAATCGATG	3005
Qy	2917	ACTGTGAGGAATGAGCAGTTCTGGTCAATTGGAGTGTCTCCACACCTCTTTGCTG	2976
Db	3006	ATTGCTGAGAAACGACGAGTTTGGGTATCGGAGGGCCCTCCGCATCTATTGCTC	3065
Qy	2977	TGTTCCAGGACTTCTCAAGGTCATAGCTGTGTGTTGATACAAGCTTCACCGTGACATCAA	3036
Db	3066	TGTTTCAAGGTTTGCTCAAAAGTTCTAGCCGAGTTTACAAGATTTCACAGTCACTTCAA	3125
Qy	3037	AGGGTGAGATGATGAGGAGTCTCAGAGCTATATACATTCAAATGGACTACCTTATTGA	3096
Db	3126	AAGCAGCAGACGATGAGGCTTCTCTGAGCTTTACATCTTCAAGTGACCAACTTTGTTGA	3185
Qy	3097	TACCTCCTACCACTTGCTTCTATTTGAACCTTCATGTTGTGCTCGCTGGCGTTCAATG	3156
Db	3186	TTCTCCGACAAACACTTCTGATCATTTACATCATTTGAGTTATTGTGCGGCTTCTGATG	3245
Qy	3157	CGATCAATTAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTATTCTTTGCAATTT	3216
Db	3246	CCATTAGCAATGGCTATGACTCATGGGAGCTCTCTTTGGGAGACTTTCTCGCTCTTT	3305
Qy	3217	GGGTGATTTGCCATCTTTATCCCTTTCTCAAAAGGTTTGTGTTGGAAGCAAAACAGACAC	3276
Db	3306	GGGTCAATTGTTCAATTATATCCCATTCCTCAAGGGAATGCTTTGGGAAGCAAAATGTC	3365
Qy	3277	CAACGATTTGCATCGTCTGTGCTTCTTCTGCTGGCTTCAATCTTCTGCTCTTTGGGTTT	3336
Db	3366	CTACGATTTATGTTGTGCTGTCTATTCTTCTAGCTTGCATCTTGAACACTCTTGTGTGGTCA	3425
Qy	3337	GGATTGATCCTTCTCTTGGCAAGGATGATGTCCTCGCTTCTTGAAGAGTGTGCTTGCATT	3396

Db 3426 GAATTAACCCGTTGTGGCTA--AAGGGGACCACTGTTGAGATCTGTGTCTGAATT 3482

QY 3397 G 3397

Db 3483 G 3483

RESULT 12

AF200526 3725 bp mRNA linear PLN 31-AUG-2000

LOCUS AF200526

DEFINITION Zea mays cellulose synthase-2 (Cesa-2) mRNA, complete cds.

ACCESSION AF200526

VERSION AF200526.1 GI:9622875

KEYWORDS

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 3725)

AUTHORS Holland, N., Holland, D., Helentjaris, T., Dhugra, K.S.,

TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene family

JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)

MEDLINE 20398328

PUBMED 10938350

REFERENCE 2 (bases 1 to 3725)

AUTHORS Dhugra, K.S. and Helentjaris, T.G.

TITLE Direct Submission

JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA

FEATURES

Source location/Qualifiers

1. .3725

/organism="Zea mays"

/mol\_type="mRNA"

/db\_xref="taxon:4577"

/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"

gene 1. .3725

/gene="Cesa-2"

CDS 179. .3403

/gene="Cesa-2"

/note="Cesa-2"

/codon\_start=1

/product="cellulose synthase-2"

/protein\_id="AAF89962.1"

/db\_xref="GI:9622876"

/translation="MAANKGMVAGSHNRNEFVMIRHGDAPVPAKPTKSANGOVCOIC

GTGVSATGDFVACNECAFPVPCYEYERKEGNOCCPOCKTRYKROKSPRVHGD

DEEDVDLDNEFNKQNGKPEWQDADLSSARHDPHRIPLTSGQOISGE

IPDASPDHSIRSPSSYVDPSVPVRIVDPSKDLNSYGLNSVDWKEVESWRVKD

KNMLQVTNKYPEARGDMEGTSGNGEDQMVDARLPLSRIVPISNQLNLYRIVILR

LILCFPPQYRISHPVNAYGLWLSVCEWPAFSLWLDQFPKMYPINRETYLDRLA

LRVDRGEPSQLAPIDVFSTVDPLKEPPLITANTVLSILAVDPYDKVSCYVSDGS

AMLTPELSLSEAFARKWVPFCCKNIEBPAPFYFAQKIDYLDKIQPSFVKERRAM

KREVEBFKIRINALVAKAOKVPEEGMTADGTAMPGNPRDHPGMIQVFLGHSGLDT

DGNELPRLVYVREKRPFGQHKHAKAGANALIRVSAVLTNQAYLNVDCDHFNSSKA

LRBAMCFMDPALGRKTCYVQFPQDFGIDLHRYANRNIVFIDINMKGLDGIQPVY

VGTGCCFNQALYGYDPLTEADLEPNIVVKSCCGRKRKRKNKSYMDSQSRIMKTES

APIFMEDIIEGIEGYEDERSVLSQKLEKRFQSPFIASFTMTQGGIPSTNPAS

LKKEAIVHISCGYEDKTEWGEIGWYGSVTEIDILTGFKHARGNOSIYCMPPRCFK

GSAPINLSRLNQVLRWALGSVEILSRHCPIMWYNGRLKLERLAYINTIYVPTS

VPLIAYCVPAICLITNKFIPEISNAGMEFILLFASIFATGILELRWSGVGIEDMW

RNEQFWVIGGSAHLFAVFOGLKVLAGIDTNTFTVTSKASDEGDFAEILYFKWTSLL

IPPTVLVNLVGMVAGISYAINSGYSGPLFGKLFISIVILHLYPFLKGLMGRQN

RTPTIVIVMSILLASIFSLWVKIDPFIISPTQKAAALGQCGVNC"

BASE COUNT 927 a 807 c 950 g 1041 t

ORIGIN

Query Match 35.0%; Score 1378.4; DB 8; Length 3725;

Best Local Similarity 66.4%; Pred. No. 3.6e-256;

Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9;

QY 119 CGCGGAGATGAGGGGAGAGAGCGATGAGCGAGCGCGGCTGTGGCCGCTCCA 178

Db 154 CGGTGTGTAGAGAGAGCGCGGAGATGGCGCCACAGAGGGATGTGGCAGGCTTCA 213

QY 179 CAACCGCAACGAGCTCGTCTCATCCCGCCGACCGGCA-----TCCCGGGCCGAAGCC 232

Db 214 CAACCGCAACGAGTTCGTATGATCCGCCACGACGCGCGCCGCTGTCCCGCTAAGCC 273

QY 233 GCCCGGGGACGAGAACGGGACGATGAGATTTGGCGGACGACGCTGGCTTGGCCCC 292

Db 274 CACGAAGAGTGCAGATGGGACAGTCTGACGATTTGTGGCGACACTGTGGCGTTTACGC 333

QY 293 CGCGGGGACCCCTTCTGTGCGTCAACGAGTGCCTTCCCGTCTGCCGAGCTGCTA 352

Db 334 CACTGTGTATGTTCTTGTGCTGCAATGATGTGCTTCCCTGTCTGCCCGCTTGTCTA 393

QY 353 CGAATACGAGCGCGGAGGACGAGAACTGCCCCAGTGCAGACTCGATACAAGCG 412

Db 394 TGAGTACGAGCGGCAAGAGGAGAACCAATGCTGCCCTCACTGACAGACTAGATACAAG 453

QY 413 CCTCAAGGCTGCCCACTGTGACCGGTGACGAGGAGGACGCGCTGATGACCTGGA 472

Db 454 ACAGAAAGGTAGCCCTCGAGTTCATGCTGATGATGAGGAGAGATGTTGATGACCTGGA 513

QY 473 CAACGAGTTCACCTGGACCGGACATGCTCGACGCTGTGTGGCGGAGTCCATGCTTACCG 532

Db 514 CAATGAATTCACCTATA-----GCAAGGCAATGGGAAGGG 549

QY 533 CCACATGAGTACGCGCGTGAAGTGAACCTATGCGCGCGCCCAAGCTTCCAGCTCAA 592

Db 550 CCCAGATGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCACTGCTCGCCATGACCC 609

QY 593 CCCCAAT--GTCCACTCCTCACCAAGGCAATAGTGTGATGACATCCACCGAGACA 649

Db 610 ACACATCGAGTTCACGCTTACAGTGAACAAGATATCTGAGAGATCCCTGATGC 669

QY 650 GCACGCGTGTGCTCTTCTTCAATGGTGTGGGGAAGAGATACATCCCTTCTTCA 709

Db 670 ATCCCTGACCGTATTTCTATCCGCACT-----CCAACATCGAGCTA 711

QY 710 TGGGATCCAGCTTACCTGTGCAACCCAGTCTATGAGCCCATCCAGGATCTTGCTGC 769

Db 712 TGTGATCAAGCGTTCAGTTCCTGTGAGGATGTGGAACCTCGAAGACTTGAAATTC 771

QY 770 ATATGGTATGTAGTGTGCTTGAAGAAAGGATGAGAAATGGAAGACAGACAAGA 829

Db 772 CTATGGGCTAATAGTGTGACTGGAAGAAAGATGAGAGCTGAGGGTTAAACAGA 831

QY 830 GAGATGACACGAGCGGGAATGATGTGTGTGATGATGTC----- 873

Db 832 CAATAATATGTTGCAAGTGAATTAATATCCAGAGGCTAGAGACATGAGAGGGAC 891

QY 874 -----ACGATGCTGATCTACCACTAATGATGAAGCAAGCAACAATGTCAGGA 925

Db 892 TGGCTCAATGAGAGATATGCAATGATGATGATGATGATGATGATGATGATGATGAT 951

QY 926 AATTCACCTTCATCAAGCCAGATTAATCATATAGATGATTAATTAATTCGGCTTGT 985

Db 952 TGTGCCAATTTCTCAACACGACTCACTTACCGGATAGTAATCATTTCTCGTCTTAT 1011

QY 986 GGTTTGGGTTCTTCTTCCACTACGAGTATGATCCGGTGAATGATGATGATGATGAT 1045

Db 1012 CATCTGTGCTTCTTCTTCCAAATATGATGATGATGATGATGATGATGATGATGATGAT 1071

QY 1046 GTGCTCATATCTGTTATCTGGAATCTGTTGCAATGCTTGAATCTTGAATCAATT 1105

Db 1072 GTGCTAGTATCTGTTATCTGAGGCTGTGCTTGTCTGCTTGTCTGATGATGATGAT 1131

QY 1106 CCCAAGTGTCTCCCTATGAGAGAGACTTACCTAGACCGGCTGCTACGTGAGCTTCA 1165

Db 1132 CCCAAATGATATCCAATCAACCGTGAGACATATCTGCACAGGCTTGCAATTGAGTATGA 1191  
QY 1166 CAAGGAAGGCCAGCCATCTCAACTTGCTCCCAATTGATTTCTTGTCAAGTACGGTTGATCC 1225  
Db 1192 TAGAGAGGAGAGCCATCAAGCTGGCTCCCATGTGATGCTTGTCAAGTACAGTGGAATCC 1251  
QY 1226 CTTAAAGGAACCTCCTTTGGTCAACAACAATACTGTTCTATCTATCTCTTCCGTTGGAATTA 1285  
Db 1252 ATTGAAGGAACCTCCACTGATCAACGCCAACACTGTTTGTCCATTCTTGCTGTGGATTA 1311  
QY 1286 TCCTGTGATAGGTTTCTTGCTATGTTTCTGATGATGCTGCTCAATGCTAACGTTTGA 1345  
Db 1312 CCTGTGTACAAAGTGTCAATGCTATGTTTCTGATGATGCTCAAGCTATGCTGACTTTTGA 1371  
QY 1346 AGCATTATCTGAAACATCTGAATTGCAAGAAGATGGTTCCTTCTGCAAAACGTAACA 1405  
Db 1372 GTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGTTCCTTTTGTAAAGACACA 1431  
QY 1406 TATTTGAACCTCGCGCTCCAGAGTGTACTTCCAACAGAGATAGACTACTTGAAGACAA 1465  
Db 1432 TATTTGAACCAAGAGCTCCAGATTTTACTTTGCTCAAAAATAGATTACCTGAAGACAA 1491  
QY 1466 GGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGACAGAGATAGAGAAATTTCAA 1525  
Db 1492 AATTCAACCTTCATTTGTTAAGAAAGACGACGAATGACAGAGATAGAAATTTCAA 1551  
QY 1526 GGTGAGATCAATGCCCTTAGTTGCCAAAGCCCAAGATTCCTGAAGAGATGACAAAT 1585  
Db 1552 AATTAAGAAATCAATGCCCTTGTTCCAAAGCACAGAAAGTGCCTGAAGAGGGGTGACCAT 1611  
QY 1586 GCAAGATGGAACCCCTGGCTGGAACAATGTTGCTGATCATCCTGGAATGATTCAGGT 1645  
Db 1612 GGCTGATGGAAGTCTTGCGCTGGGAATTAACCTAGGACCATCCTGCATGATTCAGGT 1671  
QY 1646 CTTCCTTGCCAAAGCGAGGCTTGAAGTGTGAGGGAATGAATGACGACGATGGTTTA 1705  
Db 1672 GTTCTTGGGGCACAGTGGGCTTGACACTGATGAAATGAATTAACACGCTCTGTCTA 1731  
QY 1706 TGTTTCTAGAGAGAAACGACGAGCTATTAACCATCATAGAAAGCTGGTCTATGAATGC 1765  
Db 1732 TGCTCTCGTGAAGAAAGACAGAGCTTTCAGCATCACAGAAAGCTGGTCAATGAATGC 1791  
QY 1766 ATTGGTCCGAGTCTCTGCTGACTTAACAATGCTCCATATTTGTTAACTTGGATTTGA 1825  
Db 1792 ACTGATTCGTATCTGCTGTGCTGACAAATGCTGCTATCTTCAATGTGATTTGA 1851  
QY 1826 TCACTACATCAACAACAGACAGGCTATTAAGGAAGCAATGTTTATGATGACCTTT 1885  
Db 1852 CCATTACTTCAATAGCAGCAAGCTCTTAGAAGCAATGCTTCATGATGATTCAGC 1911  
QY 1886 ACTAGGAAAGAGGTTTGTATGTACAGTTCCTCAAGATTGATGGATGATGACCA 1945  
Db 1912 TCTAGGAAGAAACTGTATGTACAAATTTCCACAAAGATTGATGGCATTGACTTGCA 1971  
QY 1946 TGACCGATATGCTAACCGGAATGTTGCTTTTGTGATATCAACATGAAGGTTGATGG 2005  
Db 1972 CGATCGATATGCTAATAGGAACATAGTCTTCTTGATATCAACATGAAGGCTTAGATGG 2031  
QY 2006 TATTTCAAGGTCCAATTTATGTGTAATCTGATGTGATTTAGAAAGCAGGCAATATATGG 2065  
Db 2032 CATTCAGGGTCCAGTCTATGTGGGAACAGGATGCTGTTCAATAGGCAGGCTTTGTATGG 2091  
QY 2066 TTATGATGCCCCCAAAACAAGAGCCACCATCAAGACTTGCACTGCTGGCCCAAGTG 2125  
Db 2092 ATAATGAT-----CCTGTTTGAAGCTGATCTGGAACCTTAACATTT 2134  
QY 2126 GTGCTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGAACTACCAAAACCAAAAC 2185  
Db 2135 GTTGTAAAG---AGCTGTGTGTAGAAAGAAAGAAAGAAAGAAAGTATATATGATAG 2190  
QY 2186 AGAGAAGAAAAAGTTATATTTTCAAGAAAGAAAGAAAGCAATCCCTGCATATGCTCT 2245  
Db 2191 TCAAAGCCGTA-----TTATGAAGAGAAACAGAAATCTTCAGCTCCCATCTTTAAT 2241

QY 2246 TGGTAATTTGACGAAGCTGCTCCAGAGCTGAGAATGAAAAAGCCGGTATTTGTAATCA 2305  
Db 2242 GGAAGACATCGAGAGGGTATTTGAAGGTTATGAGATGAAAGGTCAGTGCTTATGTCCA 2301  
QY 2306 ACAAATAATAGAAAAAATTTGGCCAACTCTGTTTGTATACATCCACTTCTCGA 2365  
Db 2302 GAGAAATTTGAGAAACGCTTTGTGAGTCTCCAACTTCATTTGATGATCCACTTTATGAC 2361  
QY 2366 GAATGTTGAACCTTGAAGAGTCAAGTCCCTGCTTCTTTTGAAGAAAGCTATACATGT 2425  
Db 2362 TCAAGGTGGCATACCACTTCAACAAACCAAGCTTCTACTGAAGGAAGCTATCCATGT 2421  
QY 2426 CATTAGTGTGTTATGAAGACAGACAGACTGGGGAAAAAGAGATTGGCTGATCTATGG 2485  
Db 2422 TATCAGCTGTGGTACGAGACAAACTGAATGGGAAAAAGAGATTGGCTGATCTATGG 2481  
QY 2486 ATCAGTTACAGAGATATTTCTAAGTGGTTCAAGATGCAATTTGTCATGTTGGCGTCAAT 2545  
Db 2482 TTCAAGTTACAGAGATATTTCTGACTGGGTTTAAATGCAATGCAAGAGCTGGCAATCAAT 2541  
QY 2546 TTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTCAGATCG 2605  
Db 2542 CTACTGCATGCCAACCAAGACCTGTTTCAAGGTTCTGCACCAATCAATCTTCTGATCG 2601  
QY 2606 TCTTCAACCAAGTCTTCGGTGGCTCTTGAGTCTATTGAGATCTTCTTCAGCAATCATTG 2665  
Db 2602 TCTTAATCAGGTGCTCCGTTGGCTCTGGGTCAGTGAAATTCGTCTAAGACATTTG 2661  
QY 2666 CCTCTTTGGTATGGGTATGTTGGCGGTGAAATTTTGAAGAAATTTCTACATCAA 2725  
Db 2662 TCCTATATGATGATGCTTCAATGAGGCGGATGAAGCTTTTGAGAGGCTGGCTTACATTAA 2721  
QY 2726 CTCCATGCTGTATCTTGGACATCTATCCCTCTTGCTTACTGTACATTCCTGCTCAT 2785  
Db 2722 CACCATTTGTTATTCGAATCACATCTGTTCCGCTTATGCGCTATTTGTGCTTCTGCTAT 2781  
QY 2786 CTGTTATTGACAGGGAATTTATACATCCAGAGCTGAATTAATGTTGCCAGCCTGTGTT 2845  
Db 2782 CTGCTTCTTACCAATTAATTTATCATCTGAGATTAGTAATTAATGTTGAATGTTCTT 2841  
QY 2846 CATGTCACTTTTATCTGCAATTTTGTCTACGACATCCCTAGAAATGAGATGAGTGTGT 2905  
Db 2842 CATTCCTTTTGTGCTCCATTTTCGCACTGGTATATTGAGCTCAGATGAGTGTGT 2901  
QY 2906 TGGAAATTGATGACTGTGGAGGAATGAGAGTCTGGGTCATTTGAGAGTGTCTCTACA 2965  
Db 2902 TGGCATTGAAGATTGGTGAAGAAATGAGCAGTTTGGGTTATTTGGTGAACCTGCCCCA 2961  
QY 2966 CCTCTTGTGTGTTCCAGGAGCTTCTCAAGTCAATAGCTGTGTGATACAAGCTTCAC 3025  
Db 2962 TCTCTTGGCGGTGTTCCAGGGTCTGCTGAAGTGTGGCTGGGATGATACCAACTTCAC 3021  
QY 3026 CGTGACATCAAGG---GTGAGATGATGAGAGTCTCAGAGCTATATACATTCAAAATG 3082  
Db 3022 AGTTACCTCAAGGCAATCTGATGAGAGTGGCACTTGTGAGCTATATGTGTCAAGTG 3081  
QY 3083 GACTACCTTATGATACCTCCTACCACTTGCTTCTATTTGAACCTTCATGTTGTGTGCTGC 3142  
Db 3082 GACCAGTTTGTCTATCCCTCCGACCACCTGTTCTGTCAATTAACCTGTCGAATGTGCGC 3141  
QY 3143 TGGCGTTTCAATGCGATCAATTAACGGAATGAGTCATGGGGCCCCCTTTGGGAAGCT 3202  
Db 3142 AGGAATTTCTGATGCAATTAACAGCGGCTACCAATCCTGGGCTCCGCTTTGGAAAGCT 3201  
QY 3203 ATTCTTTGCATTTGGGTGATGTCCATCTTTATCCCTTTCTCAAAAGTTTGGTTGAAG 3262  
Db 3202 GTTCTTCTGATCTGGGTGATCTCCATCTCTAACCCCTTCTCAAGGCTCATGGGAG 3261  
QY 3263 GCAAAACAGGACACCAACGATTTGTCATGCTGTGTCATTTCTGCTGGCTTCAATCTTCTC 3322  
Db 3262 GCAGAACCGCACGCCAACATGTCATGTTGTGTCATCTCTGCTGCTATCTTCTC 3321



QY 3323 GCTCCTTTGGGTTGCGATTGATCCTTTCT 3352  
Db 3322 CTGTGCTGTGGTGAAGATCGATCCTTTCA 3351

RESULT 13  
AF200525 3752 bp mRNA linear PLN 31-AUG-2000  
LOCUS Zea mays cellulose synthase-1 (Cesa-1) mRNA, complete cds.  
DEFINITION AF200525  
ACCESSION AF200525  
VERSION AF200525.1 GI:9622873  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 3752)  
AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,  
Xocoostle-Cazares,B. and Delmer,D.P.  
A comparative analysis of the plant cellulose synthase (Cesa) gene  
family  
TITLE Plant Physiol. 123 (4), 1313-1324 (2000)  
JOURNAL MEDLINE 20398328  
PUBMED 10938350  
REFERENCE 2 (bases 1 to 3752)  
AUTHORS Dhugga,K.S. and Helentjaris,T.G.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred  
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA  
FEATURES  
source Location/Qualifiers  
1..3752  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/note="identified and isolated from the maize genome  
database cDNA collection of Pioneer Hi-Bred International,  
Inc"  
gene 1..3752  
/gene="Cesa-1"  
183..3410  
/gene="Cesa-1"  
/note="CESA-1"  
/codon\_start=1  
/product="cellulose synthase-1"  
/protein\_id="AAF89961.1"  
/db\_xref="GI:9622874"  
/translation="MAANKGVAVGSHNRNEFMIRHGDVPGSAKPTKSANGQVCOIC  
GDSVGSATGDFVACNECAFVPCRCYERKEGNOCCPOCKTRYKROKSPRVAGD  
EDEEDVDLDNEFNFKQSGSGEMVLOQDDADLSSARHBPRIPLTSGQISGE  
IPDASPDHRHSIRPTSSYVDSVPVPRIVDPSKDLNSYGLNSVDKEREVESMRKQD  
KNMQVTNKYPEARGDMEGTSGEMOMVDARLPLSRIVPISNOLNTYRVVIL  
RLILCFPQYRVSHPRDAYGLMLVSVICEWFAISMLLDQFPKWPINRETYIDL  
ALRYDREGSPQLAPIDVFVSTVDPLKEPPLITANTVLSLSDYVDKVSQVSDG  
SAMLTSESLSAEFARKWVPCCKNIEPRAPFEYFAQKIDYLKDKIQPSFVGERA  
MKREYEEFKRINALVAKAQVPEEGMTMADGTAMPGNPRDHPGMIQVFLHSGGLD  
TDGNELRLVTVSREKRPFGFHKHKGAMNALIRVSAVLNNGAYLINDCDHYFNSSK  
ALREAMCFMMDPALGRKTYQVFPQRFDGLDHRVARNIVFPDIMKGLDGIQGPV  
YVGTGCCFNROALGYDPLVLEADLEPNIVIKSCGGRKKKNSYMDOSRIMKRTES  
SAPIFMEDIEEGIEGVEDERSVLMSORLEKRFQSPFIFASTFMTOGGIPSTNPA  
SLKEAIIHVISCGYEDKTEGKEIGMIVGSVTEIDLTFKMHARGWQSIYCMPPRCF  
KGSAPINLSRLNQLRWALGSEVILSRHCPWYGNGRLKLERLAYINTIYVPIIT  
SIPLIAYCVLPALICLTNKFIIPEISNAGMFILFASIFATGILELRMSGVIEDW  
WRNEQFVWVIGTSALHFAVFGSLKVLGIDITNFTVTSKASDEGDFAELVYFKWTSL  
LIPPTVLVINLVGMAGISAINSGYQSGPLFGKLFPSIWLHLVYFLKGLMGRO  
NRPPIVIVWSILASIFSLWVKIDPFISPTQKAALGQCGVNC"

BASE COUNT 944 a 804 c 965 g 1038 t 1 others  
ORIGIN

Query Match 34.7%; Score 1364; DB 8; Length 3752;  
Best Local Similarity 66.1%; Pred. No. 2.2e-253;  
Matches 2165; Conservative 0; Mismatches 1001; Indels 111; Gaps 9;

QY 115 TCGGCGCGGAAGTGGAGGGAGAGACCGATGAGGCGGAGCCGGCTGTGCGCGCT 174  
Db 154 TCGGTGGAGACGACGGGGAGGGAGCGGAGATGGCGGCCAACAGGGGATGTGGCGGCT 213  
QY 175 CCCACAACCGCAACGAGCTCGTGTCTATCCGCCCGCGGCGGAT-----CCCGGGCGGA 228  
Db 214 CGCACAAACCGCAACGAGTTGTCATGATCCGCCACGACGCGGATGTGCGGCTCGCTA 273  
QY 229 AGCCGCGCGGGAGCAGAACGGGACAGTGTGCGCAGATTGCGGCGACGACGTGCGCTTG 288  
Db 274 AGCCCAAAAGAGTGCAGATGACAGTCTGCCAGATTGCGGCTGACTCTGTGGGTGTT 333  
QY 289 CCCCAGCGGGGACCCCTTGTGCGTGCACACGAGTGGCGCTTCCCGTCTGCCGGA 348  
Db 334 CAGCCACTGGTATGTCTTGTGCTTCCATGATGATGTGCTTCCCTGTGCGGCCAT 393  
QY 349 GCTACGAATACGAGCGCGGGAGGACGACGACGACGACGACGACGACGACGACGACGAC 408  
Db 394 GCTATGATATGAGCGGCAAGAGGGGAACCAATGCTGCCCCAGTGCAAGACTAGATACA 453  
QY 409 AGCGCTCAAGGCGTCCCAAGTGTGACCGGTGACGAGAGAGAGAGAGAGAGAGAGAGAG 468  
Db 454 AGAGACAGAAAGTAGCCCTGAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 513  
QY 469 TGGACAACGATTCACACTGGAGCGGCATGACTCGCAGTCTGTGCGGAGTCCATGCTCT 528  
Db 514 TAGACAATGAATTCACACTACA-----GCAAGCAGTGGGA 549  
QY 529 AGCGCACATGAGTACGCGCGTGAAGTGAACCTTAATGCGCGCCACAGCTTTC--- 585  
Db 550 AAGGCCAAGAGTGGCACTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 609  
QY 586 AGCTCAACCCCAATGTTCCACTCTCTCAACCAAGGGCAATGTGATGATGATGATGATGATG 645  
Db 610 AGCCACATCATCGGATTCACGCTGACCAAGCGGTCAACAGATATCTGAGAGATTCCTG 669  
QY 646 AGCAGACGCGCTGTGCTTCTTTCATGAGTGTGGGAAAGAGATACATCCCTTC 705  
Db 670 ATGCTTCCCTGACCGCTCATCTATCCGAGT-----CCAACATGCA 711  
QY 706 CTATGCGGATCCAGCTTACTGTGCAACCCAGGTCTATGACCCATCAAGATCTTG 765  
Db 712 GCTATGTGATCCAGCGTCCAGTCTCTGTGAGAGATGTGGAACCCCTGAAAGACTTGA 771  
QY 766 CTGCATATGGTATGATGATGCTTGTGGAAGAAAGATGGAATTTGGAAGCAGAGAC 825  
Db 772 ATTCCTATGGGCTTATATGTTGACTGGAAGAAAGATTTGAGAGCTGAGGGTTAAAC 831  
QY 826 AAGAGACGATGCACGACGCGGAATGATGCTGTGTGATGATGCTG----- 873  
Db 832 AGGACAAAATATGATGCAAGTGAATAATATCAAGGCTAGAGAGAGACATGG 891  
QY 874 -----ACGATGCTGATCTACCACTAATGATGAAGCAAGACAACAACTGT 918  
Db 892 AGGGACTGGCTCAATGAGAAANATATGCAATGTTGATGATGACGCTACCTTGA 951  
QY 919 CCAAGAAAATTCCACTTCCATCAAGCCAGATTAAATCAATAGATGATATCATTAATC 978  
Db 952 GCCGATCGTCCCAATTTCTCAACAGCTCAACCTTACCGGGTAGTATCATTTCC 1011  
QY 979 GCGTTGTGTTTGGGTTCTTCTTCCACTACCGAGTATGATCCGGTGAATGATGAT 1038  
Db 1012 GTCTTATCATCTGTGCTTCTTCTTCCAGTATCGTGTCAATCATCCAGTCCGTGATCTT 1071  
QY 1039 TTGCTTGTGCTCATATCTGTTATCTGTGAATCTGTTTGGCATGCTTGAATCTTG 1098  
Db 1072 ATGATATATGCTAGTATCTGTTATCTGCGAGGCTGTGTTGCTTGTGCTTCTAG 1131  
QY 1099 ATCAATTCACAAAGTGTCTCTATTGAGAGAGACTTACCTAGACCGGCTGTCACTGA 1158  
Db 1132 ATCAGTTCCCAAATGTTATCAATCAACCGTGAAGACATATCTTGACAGGCTTGACATTTGA 1191  
QY 1159 GGTTCAGACAAAGGACGACCATCTCAACTGTCTCAATGATTTCTTGTCACTACGG 1218

Db 1192 GGTATGATAGAGAGAGAGCCATCACAGCTGGCTCCCATGTGATGTCTTCAGTACAG 1251  
QY 1219 TTGATCCCTTAAGAACTCTCTTGTGCACAACAATACTGTTCTATCTACTTTCGG 1278  
Db 1252 TGGATCCATTGAAGGAACCTCCACTGATCACAGCCAACACTGTTTGTCCATTCTTCTG 1311  
QY 1279 TGGATTATCCTGTTGATGAAGTTTCTGTATGTTTCTGATGATGCTGCAATGCTAA 1338  
Db 1312 TGGATTACCTGTGACAAAGTGCATGCTATGTTCTGATGATGCTTCAGCTATGCTGA 1371  
QY 1339 CGTTGAAGCATTATCTGAACAATCTGAATTTGCAAGAATGGTTCCTTCTGCAAC 1398  
Db 1372 CTTTGAAGTCTCTCTGAGAAACCGCAGAAATTTGTAAGAAAGTGGTCCCTTTTGTAGA 1431  
QY 1399 GGTACAATATTGAACCTCGCGCTCCAGAGTGGTACTTCCACAGAAAGATAGACTACTTGA 1458  
Db 1432 AGCACAAATATTGAACCAAGAGCTCCAGAAATTTACTTTGCTCAAAAATAGATTACCTGA 1491  
QY 1459 AAGACAAGGTGGCAGCAAACTTTGTTAGGAGAGAGCAATGAAGAGAGATATAGG 1518  
Db 1492 AGACAAAATTCAACTTCAATTTGTTAAGAAAGACGCGCAATGAAGAGGAGTATAGAG 1551  
QY 1519 AATTCAAGGTGAGAAATCAATGCTTAGTTGCAAAAGCCAGAAAGTTCCTGAAGAAGAT 1578  
Db 1552 AATTCAAAAGTAAGAAATCAATGCCCTTGTGCAAAAGCACAGAAAGTGCCTGAAGAGGGT 1611  
QY 1579 GGACAATGCAAGATGAACCCCTGGCCTGGAACAATGTTCTGATCATCTGGAATGA 1638  
Db 1612 GGACCATGGCTGATGAGAACTGCATGGCCTGGAAATATCCTAGGACCATCCTGGCATGA 1671  
QY 1639 TTCAGGTTCTTCCCTGGCCAAAGCGGAGGCTTGACTGTGAGGGAATGAATGCCACGAT 1698  
Db 1672 TTCAGGTTTCTTGGGGCACAAGTGGGCTCGACACTGATGAAATGAGTTAACACGTC 1731  
QY 1699 TGGTTATGTTTCTAGAGAGAAACGACCAAGGCTATTAACCATATGAAGAAAGCTGTCTA 1758  
Db 1732 TTGTCTATGTCTCTGTAAGAAAGAGACCAAGGCTTTCAGCATCACAAAGAGGCTGTGCAA 1791  
QY 1759 TGAATGCATTGTCGAGTCTCTGCTGTACTTAACAATGCTCCATATTTGTTAAACTTGG 1818  
Db 1792 TGAATGCGCTGATTCGTGATCTGCTGTGACAAATGGTGCTATCTTCAATGTGG 1851  
QY 1819 ATTGTGATCACTACATCAACAACAGCAAGGCTATTAAGAAAGCAATGTGTTATGATGG 1878  
Db 1852 ATTGCGACCATTAATCAATAGCAGCAAGGCTCTTAAGAAAGCAATGTGCTTCATGATGG 1911  
QY 1879 ACCCTTACTAGAAAGAGGTTTGCTATGTACAGTTCCCTCAAGATTGATGGGATTG 1938  
Db 1912 ATCCGGCTCTAGGAAGGAAACTGTTATGTACAAATTTCCACAGAGATTGATGGCATTG 1971  
QY 1939 ATGCCATGACCGATATGCTTAACCGGAATGTTGCTTTTGTGATATCAACATGAAAGGTT 1998  
Db 1972 ACTTGACGATCGATATGCTAATCGGAACATAGTTTCTTGATATCAACATGAAGGTC 2031  
QY 1999 TGGATGTTATTCAGGGTCCAATTATGTTGTAAGTGTATGTTAGAGCAGGCGAT 2058  
Db 2032 TGGATGGCATTCAAGGTCAGATTACGTGGGAACAGAGATGCTGTTCAATGACAGGCTT 2091  
QY 2059 TATATGTTATGATGCCCCCAAAACAAGAGCCACATCAAGACTTGCACTGCTGGC 2118  
Db 2092 TGTATGATACGAT-----CCTGTTTGAAGTGAAGTGAAGTGAAGCC 2134  
QY 2119 CCAAGTGTGCTTTTGTGCTGTGCTGCTTGGCAATAGGAAGCAAAAGAGACTAACCAAC 2178  
Db 2135 AAACATTGTTATTAG-----AGCTGCTGTGTAGAGAGAGAAAGAAAGACAGTTATA 2190  
QY 2179 CCAAAACAGAGAGAAAGTATATTTTTCAGAGAAAGAGAGAACCAATCCCTGCAT 2238  
Db 2191 TGGATAGTCAAAAGCCGTA-----TTATGAAGAGAAACAGAAATCTTCAGCTCCCATCT 2241  
QY 2239 ATGCTCTTGTTGAATATGAAGAGCTGCTCCAGAGCTGAAGATGAAAAGGCCGTAATTG 2298

Db 2242 TCAATATGGAAGACATCGAAGAGGTATTTGAAGTTACGAGGATGAAGGTCAGTGTCTA 2301  
QY 2299 TAAATCAACAAAATTAGAAAAGAAATTTGGCCAAATCTCTGTTTGTGTACATCCACAC 2358  
Db 2302 TGTCCAGAGGAATTTGAGAAACGCTTTGGTCAAGTCTCCCTAATTTTCATTGTCATCCACCT 2361  
QY 2359 TTCTCGAATATGTTGAACCTTGAAGAGTGCAGAGTCCCTGCTCTTTTGAAGAAGCTA 2418  
Db 2362 TTATGACACAGGTGGCATATCCACTTCAACAAACCAGCTTCTTACTAAAGAAAGCTA 2421  
QY 2419 TACATGCTATTAGTTGTGTTATGAGACAAAGACAGACTGGGAAAAAGATTGGCTGGA 2478  
Db 2422 TCCATGTCATCAGTTGTGATATGAGACAAAACCTGAATGGGAAAAAGATTGGCTGGA 2481  
QY 2479 TCTATGATCAGTTTACAGAGATATCTTAACCTGTTTCAAGATGCATTTGCTGTTGGC 2538  
Db 2482 TCTATGTTTCACTAACGAGAGATATCTGACTGGTTTAAATGCATGCAAGGGGCTGGC 2541  
QY 2539 GGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTGCAATCTTT 2598  
Db 2542 AATCAATCTACTGCATGCCACGACGACTTGTTCAGAGGTTCTGCACCAATCAATCTTT 2601  
QY 2599 CAGATCGTCTTACCAGAGTGTGCTGGGCTCTTGGGTCTATTTGAGATCTTCTTACGA 2658  
Db 2602 CCGATCGTCTTAAATCAGGTGCTCCGTTGGGCTCTTGGGTCAAGTGAATTTCTGCTTAGTA 2661  
QY 2659 ATCATTGGCCCTCTTGGTATGGGATAGGTGGCGGTCTGAATTTTGGAAAAGATTTTCT 2718  
Db 2662 GACATTGTCTATCTGTGATGGTTACAATGAGCAGATTGAAGCTTTTGGAGAGGCTGGCTT 2721  
QY 2719 ACATCAATCCATCGTGTATCTTGGACATCTATTTCCCTGTGGCTTACTGTACATTCG 2778  
Db 2722 ACATCAACACTATTGTATATCCAAATCAATCCATTCGCTTAATGCTATTTGTGCTTC 2781  
QY 2779 CTGCCATGTTTATGACAGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCC 2838  
Db 2782 CCGTATCTGCCTCTCTTACCAATTAATTTATCATTTCTGAGATTAAGCAATTATGCTGGGA 2841  
QY 2839 TGTGTTCAATGTCACCTTTTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGGA 2898  
Db 2842 TGTCTTCAATCTTCTTTTCGCTCCATTTTGGCACCTGGTATATGGAAGCTTAGATGGA 2901  
QY 2899 GTGTTGTTGAATGATGACTGTTGAGAGAAATGAGCAGTTCGGGTCAATGGAAGTGTGT 2958  
Db 2902 GTGTTGTTGCAATGAAGATTGGTGAAGAAATGAGCAGTTCGGGTATTTGGTGACACT 2961  
QY 2959 CCTCACACCTCTTGTGCTGTGTTCCAGGGACTTCTCAAGGTCAATAGCTGTGTGATACAA 3018  
Db 2962 CTGCCCATCTCTTGCAGTGTCCAGGCTGTGCTGAAGTGTGGCTGGGATTGATACCA 3021  
QY 3019 GCTTACCCGTGACATCAAAAG---GTGAGATGATGAGAGTTCACAGCTATATACAT 3075  
Db 3022 ACTTCAAGTTAACTCTCAAAAGGCACTGATGAGATGGCAGCTTGTGAGCTATATGTGT 3081  
QY 3076 TCAATGGACTACCTTATGTATACCTTACCACCTTGCTTCTATGAACTTCATTGCTG 3135  
Db 3082 TCAAGTGAACCAAGTTTGTCTCATTTCTCCGACCACTGTTCTTGTCAATTAACCTGTCGGA 3141  
QY 3136 TGTGCTGCGCTTTCAAATGCGATCAATAACGGAATATGAGTACAGGCCCCCTCTTGG 3195  
Db 3142 TGGTGGCAGAAATTTCTTATGCTCAATTAACAGTGGCTAACCAATCTGGGGTCCGCTCTTGG 3201  
QY 3196 GGAAGCTATCTTTGCAATTTTGGGTGATGTTCCATCTTTATCTTCTCAAAAGGTTGG 3255  
Db 3202 GAAAGCTGTTCTTCTGATCTGGGTGATCTTCCATCTTCACTTCCCTCAAGGCTCTCA 3261  
QY 3256 TTGAAGGCAAAACAGACACCAACGATGTCATGCTGTGCTCATCTGCTGGCTTCAA 3315  
Db 3262 TGGGAAGGAGAAACCGCACACCAACATGTCATTTGTGTTGTCATCTTCTTGCATCTA 3321  
QY 3316 TCTTCTGCTCTTTGGGTTGGAATGATCCTTCTCT 3352  
Db 3322 TCTTCTGCTGCTGTGGTGAAGATCGATCCTTCTCAT 3358

RESULT 14  
AC104487/c  
LOCUS  
DEFINITION  
AC104487 171376 bp DNA linear PLN 15-FEB-2003  
Oryza sativa chromosome 3 BAC OSJNBa0042109 genomic sequence,  
complete sequence.  
AC104487  
AC104487.3 GI:28269486  
HTG.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 171376)  
Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Gansberger,K., Jones,K.M.,  
Overton II,L.L., Tsitrin,T., Kim,M.M., Bera,J.J., Jin,S.S.,  
Padrosh,D.W., Tallon,L.J., Koo,H., Zismann,V., Hsiao,J., Blunt,S.,  
Vanaken,S.S., Riedmuller,S.B., Uterback,T.T., Feldblyum,T.V.,  
Yang,Q.Q., Haas,B.J., Suh,B.B., Peterson,J.J., Quackenbush,J.,  
White,O., Salzberg,S.L. and Fraser,C.M.  
Oryza sativa chromosome 3 BAC OSJNBa0042109 genomic sequence  
Unpublished  
2 (bases 1 to 171376)  
Buell,R.  
Direct Submission  
Submitted (12-DEC-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
3 (bases 1 to 171376)  
Buell,R.  
Direct Submission  
Submitted (07-FEB-2003) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
4 (bases 1 to 171376)  
Buell,R.  
Direct Submission  
Submitted (14-FEB-2003) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org  
5 (bases 1 to 171376)  
Buell,R.  
Direct Submission  
Submitted (15-FEB-2003) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org  
On Feb 7, 2003 this sequence version replaced gi:18497084.  
Address all correspondence to:rice@tigr.org  
COMMENT  
BAC clone OSJNBa0042109 is from Oryza sativa chromosome 3  
The orientation of the sequence is from SP6 to T7 end of the BAC  
clone.  
Genes were identified by a combination of several methods: Gene  
prediction programs including Fgenesh (<http://www.softberry.com/>),  
GENSCAN and Genscan+ (Chris Burge,  
<http://CCR-081.mit.edu/GENSCAN.html>), GenemarkHM (Mark Borodovsky,  
<http://genemark.biology.gatech.edu/Genemark/>), and Geneslicer  
(Mihaela Perlea and Steven Salzberg, contact mperlea@tigr.org),  
searches of the complete sequence against a peptide database and  
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).  
Annotated genes are named to indicate the level of evidence for  
their annotation. Genes with similarity to other proteins are named  
after the database hits. Genes without significant peptide  
similarity but with EST similarity are named as unknown proteins.  
Genes without protein or EST similarity, that are predicted by more  
than two gene prediction programs over most of their length are  
annotated as hypothetical proteins. Genes encoding tRNAs are  
predicted by tRNAscan-SE (Sean Eddy,  
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are  
identified by repeatmasker (Arian Smit,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). This BAC  
overlaps with rice BAC OSJNBa0096106 (AC092557) and OSJNBa0075M12  
(AC096856).  
FEATURES  
source  
1. 171376 Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="3"  
/map="S11301"  
/clone="OSJNBa0042109"  
/note="japonica cultivar-group"  
387.407  
/rpt\_family="AT\_rich"  
1475.1511  
/rpt\_family="AT\_rich"  
2280.6614  
/gene="OSJNBa0042109.1"  
/note="similar to Pto kinase interactor 1 GB:AA061805  
GI:3668069 (Lycopersicon esculentum); EST C72353, BF475198,  
BF475196, BE039861 from this gene"  
join(2280.2606,3008.3129,3242.4527,4612.4860,  
4958.5124,5321.5400,5486.5633,5858.6007,6091.6614)  
/gene="OSJNBa0042109.1"  
join(2280.2606,3008.3129,3242.4527,4612.4860,  
4958.5124,5321.5400,5486.5633,5858.6007,6091.6327)  
/gene="OSJNBa0042109.1"  
/codon\_start=1  
/product="putative serine/threonine kinase"  
/protein\_id="AA041118.1"  
/db\_xref="GI:28376688"  
/translation="MEGAMDGAGGVAEEGSGEEREADVAVGGGEGEEREGRT  
LVGVRAADASRTLLTWTFINVAPGDRIVAVHVLAPVAAATTAVDFTMLAYE  
GFENLKQINLKLKICKDSSVRKALVREANLFGASKVITVIGIAKKRGISLSHVAKYCS  
KLLPAKCAVLAVDSGKIVFRRESNVHSGKVAEIPGCGDNELCEVPFLAROCKEPL  
PLHEPRDGGSGAEEEEHHDVGTKETEPVNAVSGEQPSCVDPAEISPDOVGDDVP  
SDKEESTADQKDEISELPGEASVLYCVLPENHSAASTSSROSDSTEPTEGNG  
ELYCILPFRNDHSGRSSGDSRSRTASRKHDSASLSAEGDELICRLPRTGSGRSSG  
GSKRSFGAKGLIRSSSFSSDIQKDVSVCTTTEQTSMSVSTEADSPKNAARDVDP  
SSSPMSLRMIEGRPDRCLRRIFNHORSSFEWAKISMVQAMRLPSRTSVADNK  
SFKSDASPRINCDSCESTSAVDTEMSFSLYDISMPNEVESLOEKYSCTCLFTY  
ELKLATSNFSPDMLIGKGTQVYKAQLFDGLSAVKILKPSVDALIOEFVEVEIAT  
SLQHDNIVSLRGFSNDYSLVLVYDYMQLGSLDKALHGKHSKDSLWERNKIAIGI  
AKALEYLHGSVTQSVIHGDVKSNNILSEDFQQLCDFGLAKQDSASTPHLCTDIT  
GTFGYLAPEYFSGKYNKIDVYAFGVILIEISGRPIRTGCSKGQESLVGMAKPL  
SSGEIKQLVDPFLGNDYDCDEMERMTLASLCRTSSHSRPSMSOMLKLEGGDETIH  
WARTQVTASFDGSDBEAVAAPDSNMOSHNLALLGVEDDTLSHCSTEQTMDSADGYW  
SRSSSFD"  
2475.2556  
/rpt\_family="(CCG)n"  
8631.9257  
/gene="OSJNBa0042109.2"  
/note="EST CA756434 from this gene"  
8631.9257  
/gene="OSJNBa0042109.2"  
8636.9058  
/gene="OSJNBa0042109.2"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AA041120.1"  
/db\_xref="GI:28376690"  
/translation="MALDESFKRPGTIPFKMELQPGVPKQOPHHGGGGAIVSSASS  
SPAPAPAASTLLPRLPRLAPPAAASHGGCDGANILASTTTPSPSSSHRSMWSA  
RTASLVLPFTRRRGRSANSKDEDDIAFTLYGDKIV"  
9129.9159  
/rpt\_family="AT\_rich"  
9981.11794  
/gene="OSJNBa0042109.3"  
/note="predicted by fgenesh"  
join(<9981.10658,10805.10951,11258.11360,11505.>11794)  
/gene="OSJNBa0042109.3"  
join(9981.10658,10805.10951,11258.11360,11505.11794)  
/gene="OSJNBa0042109.3"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA041131.1"  
/db\_xref="GI:28376701"



```
/translation="MEFHLHARARRRKNFTIRLQDGDOLAVSHDELQELASSHFGAIL
GTSVERDFALDLHSLGLCPAHLGLGDNPFSEEEVMAVKALPSDKADGDDFTGRFYK
TCWSVIKPEVMALNSLFLANGRFSLSNDLILSLPKDEADVDERPISPISHSFG
KLFSKIPASRLSPHLGELVAPNOSASSKADPFHDFNRVQLAARALHAKHIFRLLVK
DIAKAFDTGDLSPMLFILVMDVLNMLRKASSESGFLPLNDHASLYADDLVFLSPS
VLSSIPIVHSMVIGLPAWMIKAIKKRRAPLWGNPYMDGLKAPAERAVTDMFOASTF
FVPGDESILFMSDQWIGGRSVASLADLLEVPRLRGRSRTVSSGLANNAWISDIRG
ALTPGISQFLV"
12826. .15886
/gene="OSJNBa0042109.4"
/note="EST BF475198, BF475196, BE039861 from this gene"
join(<12826. .12909,13348. .13725,14212. .14361,14548. .14627,
14713. .14866,15074. .15165,15211. .15260,15347. .15886)
/gene="OSJNBa0042109.4"
join(12826. .12909,13348. .13725,14212. .14361,14548. .14627,
14713. .14866,15074. .15165,15211. .15260,15347. .15363)
/gene="OSJNBa0042109.4"
/codon_start=1
/product="unknown protein"
/protein_id="AA041135.1"
/db_xref="GI:28376705"
/translation="MQLNPKARTRRNKTQKRVAMRDCIERDLVWRWTSRQYSAH
SAYQAFFLGHSFACVDLLLNAGPAKCFPLMFAFQRCWMADLQKRSIDSHSTCP
FCAQELETANHILDLVFPARQVLRVLSPPGNALSPVAAAGSRGTGHLPRLAAPVAG
EVFKIAGIAKALEYLNHNSITQSVIHGDVKSNNILSEDFOAQLCDFGLAKOVASAT
PRLTCTDITGTGFLAPDEYFSGKVNKIDVYAFGVILVLEIISGRHIRTGCAKGO
ESLVGMAKPLSSGGEIKQLVDLPFLGNDYDCDEMERMITLAMHKDIFSPPTGNVTDFETA
"
repeat_region 16877. .16942 /rpt_family="tandem"
repeat_region 16879. .16944 /rpt_family="CT-rich"
repeat_region 17388. .17570 /rpt_family="tandem"
repeat_region complement(18185. .18241)
repeat_region /rpt_family="Oryza"
repeat_region 18186. .18246 /rpt_family="Oryza"
repeat_region 18186. .18243 /rpt_family="Oryza"
misc_feature 18186. .18241 /note="Region with dinucleotide (AT) repeat (28-29
copies)"
repeat_region 18186. .18241 /rpt_family="tandem"
repeat_region 18332. .18414 /rpt_family="tandem"
repeat_region 18340. .18412 /rpt_family="Oryza"
repeat_region 18343. .18414 /rpt_family="Oryza"
repeat_region complement(18346. .18414)
repeat_region /rpt_family="Oryza"
repeat_region 18348. .18420 /rpt_family="Oryza"
repeat_region complement(18348. .18415)
misc_feature 18348. .18413

Query Match 33.9%; Score 1335.4; DB 8; Length 171376;
Best Local Similarity 67.5%; Pred. No. 1e-247;
Matches 2575; Conservative 0; Mismatches 456; Indels 786; Gaps 18;

QY 893 AATGATGAACAAGACAACAACACTGCCAGAAATTCACCTTCATCAAGCCAGATTAA 952
Db 139935 AATGATGAAGCTAGACAGCCATTGCTCAGAAAGATTCCGATTCTCAAGCCTGTCAA 139876

QY 953 TCCATATAGATGATTATCATTTATCGGCTTGTTGGGGTCTTCTTCCACTATCCG 1012
Db 139875 TCCCTATAGATGATTATCATATTGCGCTGTGTTCTGGGCTTCTTCTCCACTATCG 139816

QY 1013 AGTATGATCCGGTGATGATGATTTGCTTTGTGGCTCATATCTGTTATCTGTGAAT 1072
```

```
Db 139815 AGTATGATCCAGTGCCTGTATGATTCATTGCAATTATGGCTCATATCCGTGCATCTGTGAAT 139756
QY 1073 CTGTTTGCCATGCTTGGATTCTTGATTCATTCCCAAGTGGTCCCTATTGAGAGAGA 1132
Db 139755 CTGTTGCGCATGCTTGGATTCTTGATTCATTCCCAAGTGGTCCCTATTGAGAGAGA 139696
QY 1133 GACTTACCTAGACCGGCTGTCACTG----- 1157
Db 139695 AACTTACCTGACCGGTTGACATTGAGGTTTGTACTCTTGCCACCTTGTATCACTTGGC 139636
QY 1158 -----AGTTGACAAGG 1170
Db 139635 TTCTATTCTTCCGACGTTCTGTGACTTAATAATTCTTCTGATAGGTTTGACAAGG 139576
QY 1171 AAGCCAGCATCTCAACTTGTGCTCCAAATTGATTCTTTGTCAGTACGTTGATCCCTTAA 1230
Db 139575 AAGCCAGCAATCTCAACTCGCCAGTTGATTTTGTGACGTACGTTGATCCCATGA 139516
QY 1231 AGAACCTCCTTGGTCAACAACAATACTGTTCTATCTATCCTTTCGGTGATTATCCTG 1290
Db 139515 AGGAGCCTCCCTGGTCAACAACAATACTGTTCTATCTATCCTTTCGGTGATTATCCTG 139456
QY 1291 TTGATAAGGTTTCTTCTATGTTTCTGATGATGCTGCTCAATGCTTAAGTTGAAGCAT 1350
Db 139455 TCGACAAGTCTCTTCTATGTTTCTGATGATGCGGCTGCCATGCTGACATTGGAAGCAC 139396
QY 1351 TATCTGAACATCTGAATTTGCAAGAAGATGGTCTTCTTGCAACGTTACAATATTG 1410
Db 139395 TGTCTGAACATCTGAATTTGCAAGAAGTGGTCTTCTTGCAAGAAGTTACAGTCTCG 139336
QY 1411 AACCTGCGCTCCAGAGTGTACTTCCACAGAAGATAGACTTGAAGACAAGGTGG 1470
Db 139335 AGCCTGTCGCCAGAAATGTTACTTCCACAGAAGATAGACTTGAAGATAAGGTGG 139276
QY 1471 CAGCAAACTTTGTTAGGAGAGAGAGCAATGA----- 1503
Db 139275 CACCAAACTTTGTAGGAGAGAGAGCAATGAAGTACATTGCATTCTGTGTTT 139216
QY 1504 -----AGA 1506
Db 139215 GCTATGATATGAGCTCATAAATTAATGCTAATTTTGACCTCAATTCTTCTCTCAGA 139156
QY 1507 GAGAGTATGAGGAATTCAGGTGAGATCAATGCTTAGTGCCTCAAGCCAGAAAGTTC 1566
Db 139155 GCGAATATGAGGAATTCAGGTGAGATCAATGCTTAGTGCCTCAAGCCAGAAAGTTC 139096
QY 1567 CTGAAGAAGATGCAATGCAAGATGAACCCCTGGCTGGAACAATGTTGCTGATC 1626
Db 139095 CTGAGGAAGATGCAATGCAAGATGAACCTCGCTGGAAACAAGTTCGATC 139036
QY 1627 ATCCTGAATGATTC----- 1641
Db 139035 ATCCTGAATGATTCAGGTATGCTTTTACAGTTCAATGCTGTGTTAATTAATCTGCTGAC 138976
QY 1642 ----- 1641
Db 138975 ATGATGTATGCTACTTCCCTTCTCTGCTTAACCATGACAATCTCAACTTTTGC 138916
QY 1642 -----AGTCTTCTTGGCCAAAGCGAGGCTTGAAGGGAATGAATGCGCAC 1695
Db 138915 AACCATAGGCTTCTTGGCCAAAGTGTGGCCATGATGTGAAGGAATGAATGCGCTC 138856
QY 1696 GATTGTTATGTTTCTAGAGAGAAAGCAAGCAAGGCTTAATACCATCATAGAAAGCTGTG 1755
Db 138855 GATTGTTATGTTTCTAGAGAGAAAGCAAGGCTTAATACCATCATAGAAAGCTGTG 138796
QY 1756 CTATGAATGCATT----- 1768
Db 138795 CTATGAACGATTGTTAAAGACAACAGTTCAAGCATTTGTCACTTCTATAGCCTGT 138736
QY 1769 -----GCTCCAGAGTCTC 1780
Db 138735 AGCAGCATCCAGTGATATGATGTCCTTCTGTTGTCTGTTTTCAGGTCCGAGTCTC 138676
```

QY	1781	TCCTGTACTAACAAATGCTCCATATTTGTTAACTTGGATTGTGATCACTACATCAACAA	1840
Db	138675	CGCTGTACTAACTAATGCTCCATATATGTGAACTTGGATTGTGATCACTACATCAACAA	138616
QY	1841	CAGCAAGGCTATAAAGAAAGCAATGTGTTTATGATGACCCTTACTAGAAAGAGGT	1900
Db	138615	CAGCAAGGCTATAAAGAAAGCAATGTGTTTATGATGATCCCTTGTAGAAAGAAAGT	138556
QY	1901	TTGCTATGTACAGTTCCTCCAAAGATTGTATGGATTGATCCCATGACCGATATGCTAA	1960
Db	138555	TTGTTATGTGCAAGTTCCTCCAAAGATTGTATGGATTGATCCCATGATGATATGCTAA	138496
QY	1961	CCGGAATGTTGTCCTTTTGTAT-----	1982
Db	138495	TAGGAATGTCGTCCTTTTGTATGTAAGATTATCTCCAATATTTCTGATTCGTTCATCG	138436
QY	1983	-----ATCAAC	1988
Db	138435	TGTATGTTGAGCCATCAAACTTGATGCTGTTCCCTATATTAATTATTCTTGACAGATCAAC	138376
QY	1989	ATGAAGGTTTGGATGTAATCAGGGTCCAAATTTATGTTGGTACTGATGTGATTTAGA	2048
Db	138375	ATGAAGGTTTGGATGTAATCAGGGTCCCAATCTATGTCGGTACTGATGTGCTTTAGA	138316
QY	2049	AGGACGGCATTATATGTGTTATGATGCCCCCAAAACAAAGAACCCATCAAGGACTTGC	2108
Db	138315	AGACAGGCGCTGATGTTATGATGCCCCCAAAATCAAGAAACCAACATCAAGGACTTGC	138256
QY	2109	AACTGCTGCCCCAAGTGTGCTTTGCTGTGCTGCTTGGCAATAGG--AAGCAAAAG	2165
Db	138255	AACTGCTGCCCCAAGTGTGCAATTGCTGTGCTTGGTAAACAGACAAACAAAGAG	138196
QY	2166	AAGACTACCAAAACCAAAAGAGAGAAAGAAAGTTATTTTCAAGAAAGAGAGAAC	2225
Db	138195	AAGACTGCTAAACCTAAACAGAGAGAGAAAGAGGTTGTTTCAAGAGAGAGAGAT	138136
QY	2226	CAATCCCTTCATATGCTCTTGGTGAATTTGACGAAGCTGCTC-----	2268
Db	138135	CAGTCGCCCTGCATATGCTCTTGGTGAATTTGACGAAGGTCCTCAGGTATCACTTTAACT	138076
QY	2269	-----	2268
Db	138075	TCCAAGAAACCCTGTCGTTATTTAGATGCTTCCTCATCTTCATTTGATTTAGCACACT	138016
QY	2269	-----CAGAGCTGAGATGAAAAAGCCGSTATTTGTAATCAACAA	2309
Db	138015	AATCTTGAATGTTTCTGCAGAGCTGAAAAATGAAAAAGCGTGATTTGTAATCAACAG	137956
QY	2310	AAATTAGAAAAAGAAATTTGGCCAATCTTCTGTTTTTGTATACCAACACTTCTGAGAA	2369
Db	137955	AACTTAGAAAAAGAAATTTGGTCAATCTTCCGTTTTTGTGCGTCCACACTTCTGAGAA	137896
QY	2370	GGTGAACCTTGAAAGTGCAGATCTGCTTCTCTTTTGAAGAAGCTATACATGTCAT	2429
Db	137895	GGTGAACCTTGAAAGTGCAGGCCACGATCACTTTAAAGGAAGCTATACATGTCAT	137836
QY	2430	AGTTGTGTTATGAAGACAAAGACAGACTGGGGAAG-----	2466
Db	137835	AGTTGTGTTATGAAGACAAAGACAGACTGGGGAAGAAATTGAGACATGTTG	137776
QY	2467	-----	2466
Db	137775	TGTTGTTAGTTGGCAGGTGGAATATCTTCCGTTACCTTGTAAACATAACTTCTTGTCTT	137716
QY	2467	-----AGATTGGCTGGATCTATGATCAGTTACAGAAAGATATTTCTAATCTGTTCAAGAT	2521
Db	137715	TTTGTAGATTGGCTGGATCTACGATCCGTTACAGAGATATTTCTTACTGGGTTTAAGAT	137656
QY	2522	GCAATTGTCATGTTGGCGGTCAATTTAATGATACCTAAACGGGTTGATTCAAAGGTTT	2581
Db	137655	GCACTGCCATGTTGGCGGTCAATTTAATGATACCTAAAGCGGCTGCATTCAAAGGTTT	137596

QY	2582	TGACCTCTGAATCTTTGAGATCGTCTTACACGAGTCTCGGTGGGCTCTTGAGTCTAT	2641
Db	137595	TGACCTCTCAATCTTTTGACCGTCTTACACAGTCTTCCGTGGGCGCTTGATCTAT	137536
QY	2642	TGAGATCTTCTTACGAAATCATTTGCCCTCTTGGTATGGTATGGTGGCGGTCTGAATT	2701
Db	137535	TGAAATTTCTTACAGCAACCATTTGCCCTCTTGGTATGGCTATGAGGTGGATTGAATG	137476
QY	2702	TTTGAAGAATTTTCCATCATCACTCCATCGTGATCTCTTGACATCTATTCCCTCTT	2761
Db	137475	TTTGAAGAATTTTCCATCATTAATTCATCGTGATCTCTTGACATCCATTCACATATT	137416
QY	2762	GGCTTACTGTACATTGGCTCGCCATCTGTTATTGACAGGGAATTTATCACTCCAG---	2817
Db	137415	GGCTTACTGTACATTGGCTCGCCATCTGTTGCTGACAGGAATTTATCACTCCAGAGT	137356
QY	2818	-----	2817
Db	137355	AACGCTATTCTGCTCGTAATTCTGTGTGTGAATGTAGAAACCAATGACCTGAA	137296
QY	2818	-----ACGTGAATATAGTTGCCAGCCTGTGTTT	2846
Db	137295	CCTTACTTTTCTCCTCTTGTGTTCCATTCTGACAGCTTAATAATCGCAGCCTTGTGTT	137236
QY	2847	ATGTCACTTTTATCTGCAATTTTGTCTACGAGCATCTGAATATGATGAGTGTGTT	2906
Db	137235	ATGTCACTATTATCTGCAATTTTGTCTACGAGCATCTGAATATGATGAGTGTGTTA	137176
QY	2907	GGAATTGATGACTGTGTGAGGAATGAGCAGTTCTGGTCAATTGAGGTGTCTCACAC	2966
Db	137175	GGCATCATGATGATTGTGTGAGGAACGAGCAGTTTGGGTCAATTGAGGTGTCTTACAT	137116
QY	2967	CTCTTGTGTGTCTCCAGGCACTTCTCAAGTCAATAGCTGTGTGATACAAGCTTCAAC	3026
Db	137115	CTCTTGTGTGTGTCTCCAGGCACTTCTCAAGTCAATAGCTGTGTGATACAAGCTTCACT	137056
QY	3027	GTGACATCAAAAGGTGAGATGATGAGAGTTCTCAGAGCTATATACATTCAAATGACT	3086
Db	137055	GTGACATCAAAAGGTGAGATGATGAGAGTTCTCAGAGCTATATACATTCAAATGACA	136996
QY	3087	ACCTTATTGATACCTCCTACCACTTGCTTCTAATTGAATTCATTGGTGTGCTGCGC	3146
Db	136995	ACCTTATTGATACCTCCTACCACTTGCTCCTAATTGAATTCATTGGGTAAGTGTGCGC	136936
QY	3147	GTTTCAATGCGATCATATAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTATT	3206
Db	136935	GTTTCAATGCGATCATATAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTGTT	136876
QY	3207	TTTGATTTTGGGTGATGTCCATCTTATCCCTTCTCAAGGTTGGTGGAAAGCAA	3266
Db	136875	TTTGATTTTGGGTGATGTCCATCTGTATCCGTTCTGAAGGTTGGTGGAAAGCAA	136816
QY	3267	AACAGACACCAACGATGTGCATCGTGTGCCAATTCTGCTGCTTCAATCTTCGCTC	3326
Db	136815	AACAGACACCAACATGTGCATGTCTGTGCCAATTCTGCTGCTTCAATCTTCGCTC	136756
QY	3327	CTTTGGGTTCCGATTGATCTTTCTTGCGAAGGATGATGGTCCGCTTCTTGAGGAGTGT	3386
Db	136755	CTTTGGGTTCCGATTGATCTTTCTTGCGAAGGATGACGGCCCTTCTTGAGGAGTGT	136696
QY	3387	GGTTGGATTGCAACTAGATGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	3442
Db	136695	GGTTGGATTGCAACTAGAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	136636
QY	3443	GAAGTATATTTTCTGCTGTTGCCCATAT-----TTCACTGTCTGTAGAT	3488
Db	136635	GAAGTATATTTTGCAGGTGTTGTGCCCATATTTTTCATTTTTCCTCTGTAGATAG	136576
QY	3489	AAGAGACATGAATATGTCCCAAGTTTCTTTGATTCATGTGAACTTATATCTGAG	3548
Db	136575	AAACACGAAAAAACTTGCAATTTCTTTGTTCCATGTGTG--GTAGCTCAGCCTCATAT	136518
QY	3549	AGATATACTGGGGGAAAAATGAGGCTGCGGCAATCCTTGTGACGTTGGGC--CGTGAAT	3606

Db 136517 CTGCAATATTGAGCAACGGAGGCTCGACAA-CCTTGTGAGTTGCGCCACGAAAAAT 136459  
Qy 3607 ACAGCATATGCAAGTGTGATTGATGTCAGCATCTTTATTACTGTGTCGAATATAGATG 3666  
Db 136458 ACAATGAATGCCAGTGTGATTGATGTCAGCATCTTTATTAGTTGGTGAAATTGATC 136399  
Qy 3667 GGGTGAAGCCGACAGCAAGATTTGATCTGACATGCTCCGCTGACAAAC----- 3719  
Db 136398 GGGTGAAGCCGACAGCAAG-AAAAGTGTGATGATCTGATGCTCCACCATCTACAAAACT 136340  
Qy 3720 --TTGGTCTCATATAAGGACGAGATGATGCTGCC-AGTGAACAGAGCAACCTGCA 3776  
Db 136339 CGGTAATGTTCTTCTCAATAAGGACGAATGATCTGCCGCTGGAATAGAGGAAGATGCA 136280  
Qy 3777 CATTAATTTATGTA-----TGCCTGTCATTGAGGGCTGTTCATTACATGTTCTGCTA 3830  
Db 136279 CATTAATTTATGTAAGCGAGTTAAATAATGCTGAAGGGCTGTTCATTCTGTTCTCTT 136220  
Qy 3831 TACTAGAAAAACAGAAATATTAGCATTAATCTATAGTTAAATGAATGTAATGCGCC 3890  
Db 136219 TCTATTTTCATTTTTTTTGTCTGCTTTAGAACAAATAGATCTCTAGTTTGTAGTATGATGCC 136160  
Qy 3891 TGTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3927  
Db 136159 CATCCGCTGTTGTAATTTGTAAGGTGCCCTTTGTT 136123

RESULT 15  
LOCUS BT008654 3763 bp mRNA linear PLN 16-MAY-2003  
DEFINITION Arabidopsis thaliana clone RAFL09-89-G08 (R19778) putative cellulose synthase catalytic subunit (RSW1) (At4g32410) mRNA, complete cds.  
ACCESSION BT008654  
VERSION BT008654.1 GI:30794042  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 3763)  
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shin,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3763)  
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shin,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.  
The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shin,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

FEATURES  
source  
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.  
Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.  
location/Qualifiers  
1. .3763  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="RAFL09-89-G08 (R19778)"  
/note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert.  
ecotype: Columbia"  
1. .3763  
/gene="At4g32410"  
1. .144  
/gene="At4g32410"  
2  
/gene="At4g32410"  
/note="compared to genomic sequence"  
/replace="t"  
145. .3390  
/gene="At4g32410"  
/codon\_start=1  
/evidence=experimental  
/product="putative cellulose synthase catalytic subunit (RSW1)"  
/protein\_id="AAP40467.1"  
/db\_xref="GI:30794043"  
/translation="MEASAGLVAGSYRNLVLRIRHESDGTPLKMNQICQICGD  
DVGIAETGDFVACNECAFVPCRPCEYERKDGTOCCPOCKTRFRHRGSPRVEDG  
EDVDVIDENEFNYAOGANKARHGHGEEFSSSRHESQIPLLTHGTVSGEIRPDT  
QSVRTSGPLGPSDRNAISSPYIDRQVPVRIVDPKSLNSYGLGNVDMKERVEGWK  
LKQEKMLQMTGKYHEGKGGEIEGTSGNGBELQADTRLPMRVPPISSRLTPYRV  
VILRLIILCFELQYRTHPVKNAYPLMLTSVCEIWFAPSWLLDQFPKMPINRETY  
LDRLAIRYDRDEPSQLVPDVVSTVDPLKEPPLVTANTVLSLDYVPDKVACTY  
SDDGSAMLTESLSETAEFACKNVFCCKENIEBPRAPEFYAKIDYLDKIQPSFVK  
ERRAMKREYEEFKVRINALVAKOKIPEGWTMODGTPWPGNTRDHPGMIQVFLGHS  
GGLDTDGNELPRLIYVREKRPFGQHHKAGAMALIRVSAVLTNGAYLLNVDCDHYF  
NNSKAIKEAMCFMDPAIGKCCYQFPQRFQDGLDHRVANHNIVFEDIMKGLDGI  
QGPVYGTGCCFNQALGYDPLTEEDLEPNIIVKSCCGSRKKGSSKTYNEKRG  
INRSDSNAPLENMEDIDEGFEYDDERSILMSQSVKRFQSPVFIAATFMEGGIP  
PTNPATLLKEAIIHVISCGYEDKTEWKEIGWYGSVTEDLITGFMHARGWISYCN  
PPRPAFKSAPRINLSDRNLQVLRALGSEIELLSRHCPIWGYHGRLLERLRIAYINT  
IVYPITSIPLIAYCILPAFCILIDRFIIPISNYASIFILFISIAVTGILBLRWG  
VSIEDWMRNEQFWVIGTSAHLFAVFOGLKVLAVIGIDTNFTYTSKATDEDDFAELVI  
FKWTALLIPPTVLLVNLIGIVAGSVYAVNSGYQSWGPIFGKLPALMVIATLPFLK  
GLLGRQNRTPITVIVWSVLASIFSLWLRINPFDANPNANNNGKGVF"  
3391. .3763  
/gene="At4g32410"  
3' UTR  
/gene="At4g32410"  
misc\_difference 3747  
/gene="At4g32410"  
/note="not present in genomic sequence"  
BASE COUNT 997 a 768 c 892 g 1106 t  
ORIGIN  
Query Match 33.9%; Score 1332.4; DB 8; Length 3763;  
Best Local Similarity 65.0%; Pred. No. 2.8e-247;  
Matches 2117; Conservative 0; Mismatches 1066; Indels 75; Gaps 7;  
Qy 127 GTGAGGGGAGGAAGCATGAGGCGGCGGCTGTGTGCGCGGCTCCACAACCGCA 186  
Db 128 GTGTGTCGTGCTGCATGAGAGCGCAGTGCCTGTGTGCTGCATCTACCGAGAA 187  
Qy 187 ACAGCTCGTCGTCATCCGCGGACGCGCATCCCGGCGGAGCGGCGGAGCAGA 246  
Db 188 ACAGCTCGTCGTCATCCGACATGATGATGCGGCGGACCAACCTTTGAAGATATGA 247  
Qy 247 ACGGAGAGGTGTGCCAGATTGCGGCGAGCAGTCGGCCTTGCCCCGGGAGCCCT 306



Db	248	ATGGCCAGATATGTCAATCTGTGTGATGATGATGTTGCACTCGCTGAACCTGGAGATGCT	307
Qy	307	TTCGTGGCGTGCAACGAGTGCCTTCCCGCTCTGCCGGGACTGCTACGAATACGAGCGC	366
Db	308	TGTGCGGTGTAATGAATGTGCTTCCCTGTGTGTGGCTTGCTATGAGTACGAGAGGA	367
Qy	367	GGGAGGGCAGCAGAACTGCCCCAGTGCAAGACTCGATACAAGCGCCTCAAGGGCTGCC	426
Db	368	AAGATGGAAGCTAGTGTGCTTCAATGCAAGACTAGATTGACACGACAGGGGAGTTC	427
Qy	427	AACGTGTACCGGTGACGAGAGGAGGACGGCGTGCATGACTTGACAACGAGTTCAACT	486
Db	428	CTCGTGTGAAGAGATGAAGATGAGATGATGTTGATGATATCGAAGATGAGTTCA---	484
Qy	487	GGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACG	546
Db	485	-----ATTACGCCCCAGGGAGCTTAACAAGGCGAGACCAACGCCA---TGGCGAAG	532
Qy	547	GCCGTGAGGTGACCCCTTAATGCGCGGCCCAACAAGCTTTCAGCTCAACCCCAATGTTCCAC	606
Db	533	AGTTTCTTCTTCTCTAGACATGAATCTCAACCAATTCTTCTCACCCTCATGGCCATA	592
Qy	607	TCCTCACCAACGGCAAAATGCTGATGACATCCCAACGGAGACGACGCGCTGTGCTT	666
Db	593	CGGTTCTGAGAGATTGCGACGCGCTGATACACAATCTGTGCGAATAATCATCAGTCCCT	652
Qy	667	CTTTCATGGGTGTGGGGAAAGAGATATCATCCCTTCTTATGCGGATCCCAAGCTTAC	726
Db	653	-----TGGTCTTCTGACAGGAATGCTATTTCATCTCCATATATTGATCCACGGCAAC	706
Qy	727	CTGTGCAACCCAGGCTATGAGACCCATCCAAGATCTTGTGCATATGGGTATGTTAGT	786
Db	707	CTGTCCCTGAAGAATCGTGAACCCGTCAAAAGACTTGAACTTTATGGGCTTGTTAATG	766
Qy	787	TTGCTTGGAAAGAACGGAATGAGAAATTGGAAGCAGAGACAAGAGGATG-----	836
Db	767	TTGACTGGAAGAAGAGTTGAAGCTGGAAGCTGAAGCAGAGAAAAATATGTTACAGA	826
Qy	837	-----CACCAGACGGGGAATGATGCTGTGCTGATGATGCTGACGATG	879
Db	827	TGACTGCTAATACCATGAAGGAAAGGAGAGAAATTGAAGGGACTGTTCCAATGGCG	886
Qy	880	CTGATCTACCACTAATGATGAAGCAAGACAACAACCTGTCCAGAAAAATTCACCTTCAT	939
Db	887	AAGAACTCCAATGGCTGATGATACACGCTTCTATGATGCTGTGCTATCCAT	946
Qy	940	CAAGCCAGATTATCCATATAGATGATTATCATTAATCGCTTGTGGTTTGGGTTCT	999
Db	947	CTTCTGCGCTAACCCCTTATCGGGTGTGATTATTCTCGGCTTATCATCTTGTGTTCT	1006
Qy	1000	TCTTCCACTACCGAGTGAATGCCGTGAATGATGCATTTGCTTGTGGCTCATATCTG	1059
Db	1007	TCTTGCATATCGTACAACTCACCTGTGAAAAATGCATATCTTGTGTGACCTCGG	1066
Qy	1060	TTATCTGTGAAATCTGTGTTGCCATGTCTTGATTTTGAATTCATTCCCAAGTGTTC	1119
Db	1067	TTATCTGTAGATCTGTTTGCATTTCTTGCTTCTTGAATCAGTTTCCCAATGTACC	1126
Qy	1120	CTATTGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGGTTGACAAGGAAGGCCAGC	1179
Db	1127	CCATTAACAGGAGACTTATCTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAAAC	1186
Qy	1180	CATCTCAACTGTCTCCAATGATTTCTTGTCAAGTACGGTTGATCCCTTAAAGAACCTTC	1239
Db	1187	CATCACAGCTCGTTCCTGTGTGATGTGTTGTTAGTACAGTGAACCCATTGAAAGAGCTTC	1246
Qy	1240	CTTGTGTCAACAATACTGTTCTATCTATCTATCTTTCGGTGATTAATCCTGTTGATTAAG	1299
Db	1247	CCCTTGTACAGCAACACAGTTCTCTCGATTCTTCTGTGCACTAACCCGCTAGATTAAG	1306
Qy	1300	TTTCTGTATGTTCTGATGATGTGCTGCAATGCTAACGTTGAAGCATTAATCTGAAA	1359

Db	1307	TAGCCTGTTATGTTTCAGATGATGGTTCA	GCTATGCTTACCTTTGAATCCCTTCTGAAA	1366
QY	1360	CATCTGAATTTGCAAGAAATGGGTTCC	TTTCTGCAAAACGGTACAAATATTGAACCTCGCG	1419
Db	1367	CCGCTGAGTTTGCAAGAAATGGGTACCA	TTTGGCAAGAAATTCACATTTGAACCTAGGG	1426
QY	1420	CTCCAGAGTGGTACTTCCAACAGAA	GATAGTACTTGAAGACAAGGTGGCGCAAACT	1479
Db	1427	CCCCGAAATTCATTTTGGCCAGAGAT	AGATTACTTGAAGACAAAGATCCAAACCGTCTT	1486
QY	1480	TTGTTAGGGAGAGAGAGCAATGAA	GAGAGTATGAGGAATTCAGAGTGAGAAATCAATG	1539
Db	1487	TTGTTAAAGAGCGACGAGCTATGA	GAGAGGTATGAAGAGTTTAAAGTGAGATPAATG	1546
QY	1540	CCTTAGTGGCCAAAGCCAGAAAGTT	CCTGAAGAAAGATGGAACAATGCAATGGAACCC	1599
Db	1547	CTCTTGTGGCCAAAGCACAGAAAT	CCCTGAAGAGCTGGAACAATGAGATGTACTC	1606
QY	1600	CCTGGCCTGGAACAATGTTCTG	TATCATCCTGGAATGATTCAAGTCTTCTTGCCCAA	1659
Db	1607	CCTGGCCTGTAACAACACTAGAG	ATCATCTTGAATGATACAGGTCTTTAGGCCATA	1666
QY	1660	GCGAGGCTTGACTGTGAGGGAAT	GAACTGCCACGATTGGTTATGTTCTAGAGAGA	1719
Db	1667	GTTGGGGCTTGATACCGATGMA	ATGAGCTGCCACTCATGTATGTTCTCGTGAAA	1726
QY	1720	AACGACCAGGCTATTAACCATCA	TATAGAAGAGCTGCTATGAATGCAATGGTCCGAGTCT	1779
Db	1727	AGCGGCTGATTTCAACACCA	CAAAAAAGGCTGAGCTATGAATGCAATGATCCGTGTAT	1786
QY	1780	CTGCTGACTAACAATGCTCCAT	ATTTGTTAACTTGATTTGTATGATCACTACATCAACA	1839
Db	1787	CTGCTGTTCTTACCAATGAGCA	TATCTTTTGAACGTGATTTGTATGATCACTTAATA	1846
QY	1840	ACAGCAAGGCTATAAAGGAAG	CAATGTGTTTATGATGACCCCTTACTAGAGAAAGAGG	1899
Db	1847	ACAGTAAGGCTATTAAGAAG	CTATGTGTTTCAATGATGGAACCCGGCTATTGAGAAAGAGT	1906
QY	1900	TTTGCTATGTACAGTCCCTCA	AAAGATTGATGGGATTGATCGCCATGACCGATATGCTA	1959
Db	1907	GCTGCTATGTCCAGTCCCTCA	ACGTTTGAACGSTATTTGATTTGACGATGCATATGCCA	1966
QY	1960	ACCGGAATGTGTTCTTTT	TGATATCAACATGAAGGTTTGATGTATTCAGGGTCCA	2019
Db	1967	ACAGGAATATAGTCTTTT	CGATATTAACATGAAGGGGTTGATGTATCCAGGGTCCAG	2026
QY	2020	TTTATGTTGGTACTGATGT	GTATTTAGAAGGCAAGCATTATATGTTATGATGCCCCCA	2079
Db	2027	TATATGTGGTACTGGTGT	TTGTTTAATAGCGAGGCTCTATATGGGTATGCTGT	2086
QY	2080	AAACAAAGAGCCACCATCA	AGGACTTGCACTGTGCCCCAAGTGCTTTTGCTGTT	2139
Db	2087	TGACGGAAGAGATTGA	ACCAATATTA-----TTGTCAAGA	2125
QY	2140	GCTGCTTTGGCAATAGGA	AGCAAAAGAGACTACCAAAACCAAAACAGAGAAAGT	2199
Db	2126	GCTGTTCCGGGTCAAG	GAAGGTAAGTAAGCA---GAAGTATTACTACGAAAAAG	2182
QY	2200	TATTATTTTCAAGAAAG	AGAGAACCAATCCCTGCATATGCTCTTGTTGAATGACG	2259
Db	2183	GGAGAGGCATCAACAG	AGTGAATCTCACTTTCAATATGAGGACATGATG	2242
QY	2260	AAGTGCTCCAGAGCTGA	GAATGAAAAAGCCGSTATTTGTAAATCAACAAAATTAGAAA	2319
Db	2243	AGGTTTGAAGTTATG	ATGATGAGAGTCTATTTCTAATGTCCAGAGAGTGTAGAGA	2302
QY	2320	AGAAATTTGGCCAATCT	TCTGTTTGTATACCACTTCTGAGAAATGTAACCT	2379
Db	2303	AGCGTTTGGTCAGT	CGCCGSTATTTATTTGCGCAACCTTCATGGAACAAGCGGCATTTC	2362
QY	2380	TGAAGAGTGCAAGT	CTGCTTCTTTTGAAGAAAGCTATACATGTCAATTAAGTTGTGTT	2439
Db	2363	CACCAACAACCAATCC	CGCTACTCTTCTGAAAGAGGCTATTCATGTATTAAGCTGTGTT	2422

```
QY 2440 ATGAAGACAAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCAGTTACAGAAg 2499
Db 2423 ACGAAGACAAGACTGAATGGGGCAAAAGAGATTGGTTGATCTATGTTCCGTGACGGAAG 2482
QY 2500 ATATTCTAAGTGTTCAGATGATGATGTCATGTTGCGGTCAATTTACTGCAATACCTA 2559
Db 2483 ATATTCTTACTGGTTCAGATGATGATGCCCGGGTTGATATGATCTACTGCAATCCTC 2542
QY 2560 AACGGGTTCATTCAAAGGTTCTGACCTCTGAATCTTTCAGATCGTTCACCAAGTGC 2619
Db 2543 CACGCCCTGCGTTCAGGGATCTGCACCAATCAATCTTCTGATCGTTGAACCAAGTTC 2602
QY 2620 TTGGGTGGCTCTGGGTCTATTGAGATCTTCTTCAGCAATCATGCCCCCTTTGGTATG 2679
Db 2603 TTGGATGGGCTTTGGATCTATGAGATCTTCTTAGCAGACATTTGCTTATCTGSTATG 2662
QY 2680 GGTATGGTGGCGGTCTGAAATTTTGGAAAGATTTCTACATCAACTCCATCGTATC 2739
Db 2663 GTTACCATGGAAGTTGAGACTTTGGAGAGATCGTTATATCAACACCATCGTCTATC 2722
QY 2740 CTTGACATCTATTCCTCCCTCTGGCTTACTGTACATTCCTGCCATCTGTTTATGACAG 2799
Db 2723 CTATTACATCCATCCCTCTTATTTGGCTATGTATCTTCCCGCTTTTGTCTCATCACCG 2782
QY 2800 GGAATTTATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGTTCAATGTCATTTT 2859
Db 2783 ACAGATTCAATCATACCCGAGATAAGCAACTACGCGAGTATTTGTTCAATCTTCTTCA 2842
QY 2860 TCTGCATTTTGTCTACGAGCATCTAGAATGAGATGAGTGTGTTGAATTGATGACT 2919
Db 2843 TCTCAATTGCTGTGACTGAAATCTTGAGCTGAGATGAGCGGTGTGAGCATTTGAGAT 2902
QY 2920 GGTGAGGAATGAGCAGATTCTGGGTCAATTGGAGTGTCTTCAACCTCTTTGCTGTGT 2979
Db 2903 GGTGAGGAACGAGCAGATTCTGGGTCAATTGGTGACATCCGCCATCTTTTGTGTCT 2962
QY 2980 TCCAGGACTTCTCAAGGTCAATGCTGTGTGATACAAAGCTTCAACCGTACATCAAGG 3039
Db 2963 TCCAAGTCTTACTTAAGTCTTGTGTGTATCGACACCAACTTCAACCGTACATCAAGG 3022
QY 3040 ---GTGAGATGATGAGAGTCTCAGAGCTATATACATCAATGAGACTACCTTATG 3096
Db 3023 CCAAGACGAAGATGGGATTTGAGAACTCTACATCTTCAATGAGACAGCTCTCTCA 3082
QY 3097 TACCTCTTACCACCTTGTCTTATTTGAATTCATTTGTTGTTGCTGCGTTCAATG 3156
Db 3083 TTCCACCACCAACCGTCTTCTTGTGAACCTCATAGGCATTTGCTGTCTTATG 3142
QY 3157 CGATCAATTAACGATATGATGATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCAATTT 3216
Db 3143 CTGTAAACAGTGGCTACCAAGTGTGGGTCCGCTTTTCGGAAGCTCTTCTTGCCTTAT 3202
QY 3217 GGTGATTTGTCATCTTATCCCTTCTCAAGGTTTGTGGAAGCAAAACAGACAC 3276
Db 3203 GGGTATTTGCCCATCTTACCTTCTTGAAGGTCTGTGGGAAGACAAACGAAACAC 3262
QY 3277 CAAGATTGTCATGCTGTGTCATTTCTGCTGCTTCAATCTTCTGCTCTTTGGGTT 3336
Db 3263 CAACCATGTCATTTGTCTGTGTTCTTCTGCTCCATCTTCTGTTGCTTTGGGTCA 3322
QY 3337 GGATGATCTCTTCTCTG 3354
Db 3323 GGATCAATCCCTTTGTGG 3340
```

Search completed: December 15, 2003, 22:42:12  
Job time : 14024 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 17:49:46 ; Search time 7630 Seconds

(without alignments)  
12537.669 Million cell updates/sec

Title: US-09-720-383C-9

Perfect score: 3936

Sequence: 1 ctctccctcgtcgtcgcg.....ctgaattggtttgtgaaaa 3936

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_htc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_htc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rtd: \*  
26: em\_gss\_phg: \*  
27: em\_gss\_vr1: \*  
28: gb\_gss1: \*  
29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869.8	98.3	3897	11	AY103655 Zea mays
2	2591	65.8	3783	11	AY104730 Zea mays
3	2225.2	56.5	3788	11	AY103701 Zea mays
4	1342.8	34.1	3898	11	AY110415 Zea mays

5	1289.4	32.8	3763	11	AY108113 Zea mays
6	1212.4	30.8	3696	11	AY110079 Zea mays
7	1209.6	30.7	3728	11	AY112236 Zea mays
8	1161.6	29.5	2872	11	AY104236 Zea mays
9	811	20.6	877	29	CC379931 PUHOP50TB
10	731.4	18.6	758	14	CD441181 EL01N0553
11	704.6	17.9	856	14	CB633016 OS1EB11J
12	686.2	17.4	859	14	CB645001 OSJNEB06H
13	683.2	17.4	864	14	CB627056 OS1EB01L
14	672.4	17.1	862	14	CD437450 EL01N0501
15	670.6	17.0	900	29	CC365622 PUHBR52TD
16	652	16.6	792	14	CB671068 OSJNE04H
17	641.2	16.3	783	14	CB671503 OSJNE05B
18	626.6	15.9	1022	29	BZ827769 PUFH017TB
19	621.8	15.8	795	14	CB677472 OSJNE14G
20	612.8	15.6	806	14	CB629322 OS1EB05H
21	596.4	15.2	758	14	CB6292249 OSJNE06D
22	594.6	15.1	714	14	CB629928 OS1EB06I
23	580.6	14.8	657	9	AM037064 614025D05
24	578.6	14.7	765	14	CB676377 OSJNE12J
25	570.6	14.5	670	29	CC160666 ig10C05.g
26	566.6	14.4	827	14	CB677498 OSJNE14I
27	549	13.9	701	13	BQ803259 WHE2835_E
28	546.8	13.9	770	14	CB681901 OSJNEf08K
29	539.6	13.7	789	14	CB672865 OSJNE07C
30	539	13.7	591	14	CD425359 SA1_11_F0
31	537.8	13.7	930	29	CC354726 PUFH20TB
32	537.6	13.7	635	9	AM146609 614074E01
33	533.4	13.6	642	9	AM053172 614072C08
34	532	13.5	660	9	AI621823 486027A04
35	525.4	13.3	539	13	BU197567 946162C06
36	524.8	13.3	745	14	CD444028 EL01N0434
37	524.4	13.3	695	13	BU100946 WHE3359_G
38	522.6	13.3	755	13	BQ484013 WHE3515_D
39	517.2	13.1	760	10	BE412810 MCG008.D0
40	517	13.1	693	10	BF628750 HVSMEB000
41	515	13.1	700	9	AL503528 AL503528
42	510	13.0	510	9	AM126453 614072C08
43	505.4	12.8	515	9	AI920493 614025D05
44	502.8	12.8	731	13	BQ294855 WHE2855_B
45	501.8	12.7	619	14	CA019440 HV11P23x

ALIGNMENTS

RESULT 1  
AY103655  
LOCUS AY103655 3897 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PC0096398 mRNA sequence.  
ACCESSION AY103655  
VERSION AY103655.1 GI:21206733  
KEYWORDS  
SOURCE HTC.  
ORGANISM Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1 (bases 1 to 3897)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitlitt,M.S.,  
Arthur,L.W., Hanfey,M., Morgante,M. and Tingey,S.V.

TITLE  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes

JOURNAL  
Unpublished (2002)  
2 (bases 1 to 3897)  
Coe,B.H.

REFERENCE  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA

COMMENT  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSL, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iasate.edu](http://www.zmdb.iasate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the





Db 1813 CATATTTGTTAAACTTGATGTGTATCACTACATCAACAAGCAAGGCTATAAAGAAAG 1872

QY 1861 CAATGTGTTTATGATGAGCCCTTACTAGGAAAGAGTTTCTATGTACAGTTCCTC 1920

Db 1873 CAATGTGTTTATGATGAGCCCTTACTAGGAAAGAGTTTCTATGTACAGTTCCTC 1932

QY 1921 AAAGATTTGATGGGATTGATCCCATGACCATATGCTAACCGGAATGTTGCTTTTGG 1980

Db 1933 AAAGATTTGATGGGATTGATCCCATGACCATATGCTAACCGGAATGTTGCTTTTGG 1992

QY 1981 ATATCAACATGAAAGGTTTGGATGTATTCAGGGTCCAAATTATGTGTACTGATGTG 2040

Db 1993 ATATCAACATGAAAGGTTTGGATGTATTCAGGGTCCAAATTATGTGTACTGATGTG 2052

QY 2041 TATTTAGAAAGCAGGCAATTATGCTTATGATGCCCCCAAAACAAGAGCCACCATCA 2100

Db 2053 TATTTAGAAAGCAGGCAATTATGCTTATGATGCCCCCAAAACAAGAGCCACCATCA 2112

QY 2101 GGACTTGCAACTGCTGGCCCAAGTGTGCTTTTGTGCTGCTTTGGCAATAGGAAGC 2160

Db 2113 GGACTTGCAACTGCTGGCCCAAGTGTGCTTTTGTGCTGCTTTGGCAATAGGAAGC 2172

QY 2161 AAAAGAGAAGCTACCAAAACCAAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220

Db 2173 AAAAGAGAAGCTACCAAAACCAAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2232

QY 2221 AGAACCAATCCCTGCAATATGCTCTGTGAAATTGACGAAGCTGCTCCAGAGCTGAGA 2280

Db 2233 AGAACCAATCCCTGCAATATGCTCTGTGAAATTGACGAAGCTGCTCCAGAGCTGAGA 2292

QY 2281 ATGAAAAGCCCGTATTGTAAATCAACAAAATTAGAAAAGAAATTGGCCATCTTCTG 2340

Db 2293 ATGAAAAGCCCGTATTGTAAATCAACAAAATTAGAAAAGAAATTGGCCATCTTCTG 2352

QY 2341 TTTTGTACATCCACACTTCTGAGAAATGTTGGAACCTTGAAGAGTGCAGTCCCTGCTT 2400

Db 2353 TTTTGTACATCCACACTTCTGAGAAATGTTGGAACCTTGAAGAGTGCAGTCCCTGCTT 2412

QY 2401 CTCTTTTGAAGAGCTATACATGTCAATTAAGTGTGTTATGAAGACAAAGACAGACTGG 2460

Db 2413 CTCTTTTGAAGAGCTATACATGTCAATTAAGTGTGTTATGAAGACAAAGACAGACTGG 2472

QY 2461 GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAAAGATTTCTACTGTTCAAGA 2520

Db 2473 GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAAAGATTTCTACTGTTCAAGA 2532

QY 2521 TGCATTTGTCATGTTGGCGGTCAATTTACTGCATACCTAAACGGTTCATTCAAAGTT 2580

Db 2533 TGCATTTGTCATGTTGGCGGTCAATTTACTGCATACCTAAACGGTTCATTCAAAGTT 2592

QY 2581 CTGCACTCTGAAATCTTTCAAGATCGTCTTCAACCAAGTCTTGGGTGGGCTCTTGGGCTTA 2640

Db 2593 CTGCACTCTGAAATCTTTCAAGATCGTCTTCAACCAAGTCTTGGGTGGGCTCTTGGGCTTA 2652

QY 2641 TTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGTTGGCGGTCTGAAT 2700

Db 2653 TTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGTTGGCGGTCTGAAT 2712

QY 2701 TTTTGAAGAGATTTCCTACATCAACTCCATCGTATCTTGGACATCTATCCCTCT 2760

Db 2713 TTTTGAAGAGATTTCCTACATCAACTCCATCGTATCTTGGACATCTATCCCTCT 2772

QY 2761 TGGCTTACTGTACATTTGCCCTGCATCTGTTTATGACAGGAAATTATCACTCCAGAGC 2820

Db 2773 TGGCTTACTGTACATTTGCCCTGCATCTGTTTATGACAGGAAATTATCACTCCAGAGC 2832

QY 2821 TGAATAATGTGCGCAGCTGTGTTCATGTCACTTTTATCTGCATTTTGTCTACAGCA 2880

Db 2833 TGAATAATGTGCGCAGCTGTGTTCATGTCACTTTTATCTGCATTTTGTCTACAGCA 2892

QY 2881 TCCTAGAAATGAGATGAGTGTGTGTAATTGATGACTGTGAGGAATGAGAGTTCT 2940

Db 2893 TCCTAGAAATGAGATGAGTGTGTGTAATTGATGACTGTGAGGAATGAGAGTTCT 2952

QY 2941 GGGTCATTGGAGGTGTGCTCTCAACACTCTTGTGTGTTCACAGGACTTCTCAAGTCA 3000

Db 2953 GGGTCATTGGAGGTGTGCTCTCAACACTCTTGTGTGTTCACAGGACTTCTCAAGTCA 3012

QY 3001 TAGCTGTGTGATACAACTTCAACCGTGACATCAAGGGTGAGATGATGAGAGTTCT 3060

Db 3013 TAGCTGTGTGATACAACTTCAACCGTGACATCAAGGGTGAGATGATGAGAGTTCT 3072

QY 3061 CAGAGCTATATACATTCAAATGAGTACCTTATGATACTCTTACCACTTGTCTTAT 3120

Db 3073 CAGAGCTATATACATTCAAATGAGTACCTTATGATACTCTTACCACTTGTCTTAT 3132

QY 3121 TGAACCTTATGTTGTGTGCTGCTGCGCTTCAATGCGATCAATAACGGAATAGAGTCAT 3180

Db 3133 TGAACCTTATGTTGTGTGCTGCTGCGCTTCAATGCGATCAATAACGGAATAGAGTCAT 3192

QY 3181 GGGGCCCCCTCTTGGGAAGCTATTCTTTGCAATTTTGGGTGATTGTCCAATCTTATCCCT 3240

Db 3193 GGGGCCCCCTCTTGGGAAGCTATTCTTTGCAATTTTGGGTGATTGTCCAATCTTATCCCT 3252

QY 3241 TTCTCAAGGTTTGTGTTGGAAGCAAAACAGACACCAACGATTGTCTGTGTCCA 3300

Db 3253 TTCTCAAGGTTTGTGTTGGAAGCAAAACAGACACCAACGATTGTCTGTGTCCA 3312

QY 3301 TTCTGTGCTTCAATCTTCTGCTCTTGGGTTGCGATTGATCTTCTTGTGCGAAGG 3360

Db 3313 TTCTGTGCTTCAATCTTCTGCTCTTGGGTTGCGATTGATCTTCTTGTGCGAAGG 3372

QY 3361 ATGATGTCCTGCTTCTGAGAGTGTGTTGGAATTGCAACTAGAGTCAAGTGCATCAG 3420

Db 3373 ATGATGTCCTGCTTCTGAGAGTGTGTTGGAATTGCAACTAGAGTCAAGTGCATCAG 3432

QY 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGTGTGTGTCGCCATATTCAGTGT 3480

Db 3433 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGTGTGTGTCGCCATATTCAGTGT 3492

QY 3481 CTGTAGATAAGAGACATGAATATGCCAAGTTCTTTTGTATCCATGTGTGAACCTACTTAA 3540

Db 3493 CTGTAGATAAGAGACATGAATATGCCAAGTTCTTTTGTATCCATGTGTGAACCTACTTAA 3552

QY 3541 TATCTGAGAGATATATCTGGGGGAAATGAGGCTGCGCAATCTTGTGCAAGTTGGCCG 3600

Db 3553 TATCTGAGAGATATATCTGGGGGAAATGAGGCTGCGCAATCTTGTGCAAGTTGGCCG 3612

QY 3601 TGAATACAGCATATGCAAGTGTGTTGATTTGACAGCATTTCTTATTAATGTTGGTCAATA 3660

Db 3613 TGAATACAGCATATGCAAGTGTGTTGATTTGACAGCATTTCTTATTAATGTTGGTCAATA 3672

QY 3661 TAGATGGGCTGAGCCGAACAGCAAGGTAATTTGATTTCTGCACTGCTCCGTTGACAACT 3720

Db 3673 TAGATGGGCTGAGCCGAACAGCAAGGTAATTTGATTTCTGCACTGCTCCGTTGACAACT 3732

QY 3721 TGGTCTCAATTAAGCAGCAGGCAAGATGCATCTGCCAGTGGAAACAGAGCAACCTGCACATT 3780

Db 3733 TGGTCTCAATTAAGCAGCAGGCAAGATGCATCTGCCAGTGGAAACAGAGCAACCTGCACATT 3792

QY 3781 ATTATGTATGCTGTTCATTGAGGGCTTGTTCATTACATGTTGCTATTAAGTGAATA 3840

Db 3793 ATTATGTATGCTGTTCATTGAGGGCTTGTTCATTACATGTTGCTATTAAGTGAATA 3852

QY 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAGTATGTAAT 3885

Db 3853 AACAGATATTAGCATTAATCTATAGTTAATTAAGTATGTAAT 3897

RESULT 2  
AY104730 3783 bp mRNA linear HTC 16-OCT-2002  
LOCUS  
DEFINITION Zea mays PCO100501 mRNA sequence.  
ACCESSION AY104730  
VERSION AY104730.1 GI:21207808  
KEYWORDS HTC.

SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 3783)
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL	Unpublished (2002)
REFERENCE	2 (bases 1 to 3783)
AUTHORS	Coe,E.H.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT	Missouri, Columbia, MO 65211, USA If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES	Location/Qualifiers
Source	1. .3783

```

/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

Query Match	65.8%	Score 2591;	DB 11;	Length 3783;
Best Local Similarity	86.5%;	Pred. No. 0;		
Matches 2884; Conservative	0;	Mismatches 440;	Indels 9;	Gaps 2;

[illegible]

Db	566	CATGCTCCACGCGCCACATGAGCTACGCGCGCGCCGACCTCGACGGCGTGCAGCC	625
Qy	581	TTTCCAGCTCAACCCCAATGTTCCACTCTCTCAACGCGCAATAGTGTGATGACATCC	640
Db	626	ATTCCACCCCATCCCAATGTTCCTCCTCCTCAACGACGACAGATGTCATGACATCC	685
Qy	641	ACCGAGCAGCACCGCGCTGCTGCTCTTTCATGGGTGTGGGGAAAGAGATACATCC	700
Db	686	GCCGGACCAAGCACGCGCTTGTGCTCCTGTCGTGGTGGCGGGGAAGAGATTCAACC	745
Qy	701	CCTTCCTTATGCGGATCCAGCTTACCTGTGCAACCAAGTCTATGAGCCCATCAAGGA	760
Db	746	TCTCCGTACCGGATCCCACTTCTCTGTGCAACGAGGTCATGGAACCTTCCAGGA	805
Qy	761	TCTTGCTGATATGGGTATGCTAGTGTGCTTGGAAAGAACGATGGAATTGAAGCA	820
Db	806	TCTCGCCGATATGGCTACGGGAGCGTAGCATGGAAGAGAGATGAGAGCTGGAACA	865
Qy	821	GAGACAAGAGAGATGCACCAAGCGGGAATGATGCTGTGTGATGATGTTGACGATGC	880
Db	866	GAAAGCAGAGAGATGCACCAAGCAGGAACGATGGCGCGCGGATGATGTTGATGTC	925
Qy	881	TGATCTACCACTAATGATGAGCAGCAACAACCTGTCCAGGAAATTCACCTTCATC	940
Db	926	AGATCTACCACTAATGATGAGCAGCAGCAACCTGTCTCAGAAAGATCCCGCTTCTC	985
Qy	941	AAGCCAGATTAACTCATATAGGATGATTATCATTAATTCGGCTTGTGTTGGGTTCTT	1000
Db	986	AAGCCAATCAACCCCTATAGGATGATTATAATAATTCGCTAGTGTGTTGTCTTCTT	1045
Qy	1001	CTTCCACTACCGAGTGATGCATCCGGTAATGATGCAATTGCTTGTGCTCATATCTGT	1060
Db	1046	CTTCCACTACCGAGTGATGCATCCGGTGCTGATGATGCTTATATGCTCATATCTGT	1105
Qy	1061	TATCTGTAATCTGGTTGCTGCATGTCTTGATCTTGATCAATTCCAAAGTGTCCC	1120
Db	1106	GATCTGTAATTTGGTTGGCATGTCTTGATCTTGACCAAGTTCCAAAGTGTCCC	1165
Qy	1121	TATTGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGGTTCCACAAGAAAGCCACC	1180
Db	1166	TATCGAGAGGAAACCTATCTTGACCGGCTGAGTTAAAGTTTGACAAGAAAGGCAATC	1225
Qy	1181	ATCTCAACTGCTCCAAATTGATTTCTTGTCACTACGAGTTGATCCCTTAAAGAACCTCC	1240
Db	1226	TTCTCAACTCGCCCTGTGTGATTTCTTGTGACGAGTTGATCCCTTGAAGAACCTCC	1285
Qy	1241	TTTGGTCAACAACAATACTGTCTATCTATCTTTCGCTGAGATTATCCGTGTGATAAGT	1300
Db	1286	ATTGGTCACTGCTAATACTGTCTATCTATCTTTCGCTGAGATTATCCAGTTGATAAGT	1345
Qy	1301	TTCTTGCTATGTTCTGTGATGCTGCTGCAATGCTAAAGTTTGAAGCATTATCTGAAAC	1360
Db	1346	TTCATGCTACGTTCTGTGATGCTGCTGCAATGCTAAAGTTTGAAGCATTATCTGAAAC	1405
Qy	1361	ATCTGAATTTGCABAAGAAATGGGTTCTCTTCTGCAACGCTACATATTTGAACCTCGCGC	1420
Db	1406	ATCTGAATTTGCABAAGAAATGGGTTCTCTTCTGCAAAAGATATTAAGCTTGAACCTCGTGC	1465
Qy	1421	TCCAAGATGCTACTTCCACAGAGATAGACTACTTGAAGAACAGAGTGGCAGCAAACTT	1480
Db	1466	TCCAAGATGCTACTTCCACAGAGATAGACTACTTGAAGAACAGAGTGGCGCAAACTT	1525
Qy	1481	TGTTAGGAGAGAGAGCAATGAAGAGAGATATGAGGAATTCAGGTGAGATCAATGC	1540
Db	1526	TGTTAGAGAACGAGAGCAATGAAGAGAGATATGAGGAATTCAGGTGAGATCAATGC	1585
Qy	1541	CTTAGTTGCCAAAGCCAGAAAGTCTCTGAAGAAGGATGGAACAATGCAAGATGGAACCC	1600
Db	1586	CTTGTTGCTAAAGCCCAAAAGTCTCTGAGGAAGGATGGAACAATGCAAGATGGAACCTC	1645
Qy	1601	CTGGCTGGAACAATGTTGCTGATCATCTGGAATGATTCAGGTCTTCTTGCCAAAG	1660



Db 1646 ATGGCCCGGAATATATGTCCGTGATCATCCTGGAATGATTCAGGTTTCTTGTCAAAG 1705  
Qy 1661 CGGAGGCGCTTGA CTGTGAGGGAATGA CTGCCACGATTTGGTTATGTTCTAGAGAGA 1720  
Db 1706 TGGTGGCCATGATGTGGAAGGAATGAGCTGCCCTCGATTGGTTATGTTTCAAGAGAGAAA 1765  
Qy 1721 ACGACGAGGCTATTAACCATCATAGAAAAGCTGGTGCATGATGATTTGGTCCGAGTCTC 1780  
Db 1766 ACGGCCAGGCTACAACCATCACAGAGGCTGGTGCATGATGATTTGGTCCGAGTCTC 1825  
Qy 1781 TGGTGTACTAACAATGCTCCATATTGTTAACTTGATTTGATTCATCACTCAACAA 1840  
Db 1826 TGGTGTACTAACAATGCTCCATATTGCTGAACCTTGATTTGATCACTATATCAATAA 1885  
Qy 1841 CAGCAAGGCTATAAAGGAACAATGTTTATGATGACCCCTTACTAGAAAGAGGT 1900  
Db 1886 TAGTAAGGCTATAAAGGAACAATGTTTATGATGATCCCTTGTGGAATAAAGT 1945  
Qy 1901 TTGCTATGTACAGTTCCTCCAAAGATTGTGATGGGATGATCCGCATGACCGATATGCTAA 1960  
Db 1946 TTGCTATGTGACAGTTTCTCCAAAGATTGTGATGGGATGATCCGCATGATGATATGCTAA 2005  
Qy 1961 CCGGAATGTTGCTTTTGTGATATCAATGAAGGTTTGATGTTATTCAGGCTCAAT 2020  
Db 2006 CAGAAATGTTGCTTTTGCATATCAATGAAGGTTTGATGTTATTCAGGCTCAAT 2065  
Qy 2021 TTATGTGTGACTGTGATGTATTTGAAGGAGGAGGATTTATGTTATGATGCCCCCAA 2080  
Db 2066 TTATGTGTGACTGTGATGTCTTCAAGAGGAGGATTTATGCTTACGATGCTCCCAA 2125  
Qy 2081 AACAAAGAGCCACCATCAAGGACTTGCACTGCTGGCCAAAGTGTGCTTTTGCTGTG 2140  
Db 2126 AACAAAGAGCCACCATCAAGACTTGCACTGCTGGCCAAAGTGTGCTTTTGCTGTG 2185  
Qy 2141 CTGCTTTGGCAATAGGAAGCAAAAGACTAACCAACCCAAAACAGAGA-----AGAA 2194  
Db 2186 CTGTTTGGTAACAGGAAGCAAGAAAGAACCAAGACCCTTAACCTTAATTGAGAA 2245  
Qy 2195 AAAGTTATTTTTCAGAAAGAAAGAACCAATCCCCTGCATATGCTCTTGTTGAAT 2254  
Db 2246 GATTAAGAACTTTTAAAGAAAAGAAATCAAGCCCCCTGCATATGCTCTTGTTGAAT 2305  
Qy 2255 TGACGAGCTGCTCCAGAGCTGAGATGAAAAAGCCGATTTGTAATCAACAAAAAT 2314  
Db 2306 TGATGAGCTGCTCCAGAGCTGAATAAGGCTAGTATTTGAATCAACAGAAATT 2365  
Qy 2315 AGAAAGAAATTTGGCCATCTTCTGTTTGTACATCCACACTCTCGAAGATGTTG 2374  
Db 2366 GGAAGAAATTTGGCCAGTCTTCAATTTTGTGCATCCACACTCTTGAGAATGTTG 2425  
Qy 2375 AACCTTGAAGAGTGCAGTCTGCTTCTTTTGAAAGAAAGCTATACATGTCATTAGTTG 2434  
Db 2426 AACCTTGAAGAGTGCAGTCTGCTTCTTTGTAAGGAAAGCTATACATGTCATTAGTTG 2485  
Qy 2435 TGGTTATGAAGACAAGACAGACTGGGAAAAAGATTGGCTGATCTATGATCAGTTAC 2494  
Db 2486 TGGATATGAAGACAAGACAGCTGGGAAAAAGATTGGTTGATTATGATCAGTCAAC 2545  
Qy 2495 AGAAGATATTCTAAGCTGTTCAAGATGATGTCATGTTGGCGTCAATTACTGCAT 2554  
Db 2546 AGAAGATATTCTAAGCTGTTCAAGATGATGTCATGTTGGCGTCAATTACTGCAT 2605  
Qy 2555 ACCTAAAGGGTTGCATTCAAGGTTCTGACCTCTGAATCTTTCAGATGCTTACCA 2614  
Db 2606 ACCTAAAGGGCGGCTTCAAGGTTCCGACCTCTCAATCTTTCGATCGTCTTACCA 2665  
Qy 2615 GGTGCTCGTGGGCTCTTGGGCTATTTAGATCTTCTCAGCAATCATTTGCCCTCTTG 2674  
Db 2666 GGTCTTCCGTGGGCTCTTGGTTCATTTGAATTTTCTTCAAGCAACACTGCCCTCTCTG 2725  
Qy 2675 GTATGGGTATGTTGGCGGTCTGAATTTTGGAAAAGATTTCTTACATCAACTCCATCGT 2734  
Db 2726 GTATGGGTATGTTGGTGAAGTTCTTGAAAAGGTTTTCGTACATTAACTCCATCGT 2785

Qy 2735 GTATCCTTGACATCTATTCCTCTCTTGCTTACTGATGATTCGCTGCCATCTGTTATT 2794  
Db 2786 ATACCTTGGACATCTATCCGCTCTTGCCCTATTTGACATTTGCCATCTGCTTGTCT 2845  
Qy 2795 GACAGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCTGTGTTCAATCACT 2854  
Db 2846 GACAGGAAATTTATCAAGCAGAGCTTAACAATGTTGCCAGCTGTGTTCAATCACT 2905  
Qy 2855 TTTTATCTGCATTTTGTCTACGAGCATCTCTAGAAATGATGAGTGTGTTGAATTGA 2914  
Db 2906 TTTTATCTGCATTTTGTCTACGAGCATCTCTAGAAATGATGAGTGTGTTGAATTGA 2965  
Qy 2915 TGACTGTGAGGAATGAGAGTCTTGGGTCAATTGAGGTGTCTTCAACACTCTTTTG 2974  
Db 2966 TGACTGTGAGGAATGAGAGTCTTGGGTCAATTGAGGTGTCTTCAACACTCTTTTG 3025  
Qy 2975 TGTGTTCCAGGACTTCTCAAGGTCATGAGTGTGTTGATPACAAGCTTCAACGTCATC 3034  
Db 3026 TGTGTTCCAGGACTTCTCAAGGTCATGAGTGTGTTGATPACAAGCTTCAACGTCATC 3085  
Qy 3035 AAAGGTGAGATGATGAGAGTCTCAGAGCTATATACATTCAATGAGACTTACTTATT 3094  
Db 3086 CAAGGGCGGAGACGACGAGAGTCTCAGAGCTGTACACATTCAATGAGACGACCTTCT 3145  
Qy 3095 GATACCTCCTACCACTTGTCTTATGAACTTCAATTGCTGTGCTGCGTGGCTTCAAA 3154  
Db 3146 GATACCTCCTACCACTTGTCTTATGAACTTCAATTGCTGTGCTGCGTGGCTTCAAA 3205  
Qy 3155 TGGGATCAATTAACGGATATGAGTCAATGAGGCCCCCTCTTTGGGAAGCTATTCTTGCA 3214  
Db 3206 TGGGATCAATTAACGGATATGAGTCAATGAGGCCCCCTCTTTGGGAAGCTATTCTTGCA 3265  
Qy 3215 TTGGGTGATTTGTCATCTTATTCCTTTCAAAGGTTTGGTTGGAAGCAAAACAGAC 3274  
Db 3266 TTGGGTGATTTGTCATCTTATTCCTTTCAAAGGTTTGGTTGGAAGCAAAACAGAC 3325  
Qy 3275 ACCAAGATTGTCATGCTGTGCTCATTCTGCTGCTTCAATCTTCTGCTCTTGGGT 3334  
Db 3326 GCCAAGATTGTCATGCTGTGCTCATTCTGCTGCTTCAATCTTCTGCTCTTGGGT 3385  
Qy 3335 TCGGATGATCTTCTTCCCTGGAAGGATGATGTTCCGCTTCTTGAGAGTGTGTTGGA 3394  
Db 3386 CCGGATCGACCCGCTTCTTCCGAAGGATGATGTTCCGCTTCTTGAGAGTGTGTTGGA 3445  
Qy 3395 TTGCAACTAGGATGTCAGTGCATCAGCTCCCC 3427  
Db 3446 TTGCAACTAGGAGTCAAGCTGAGCTTCCCC 3478

RESULT 3  
AY103701  
LOCUS 3788 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PCO120363 mRNA sequence.  
ACCESSION AY103701  
VERSION AY103701.1 GI:21206779  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
REFERENCE 1 (bases 1 to 3788)  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 3788)  
AUTHORS Coe,B.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA

COMMENT

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR, [www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES

Source

Location/Qualifiers  
1..3788  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:637005"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Cornsensus library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT

851 a 954 c 1114 g 866 t 3 others

ORIGIN

Query Match 56.5%; Score 2225.2; DB 11; Length 3788;

Best Local Similarity 81.1%; Pred. No. 0;

Matches 2685; Conservative 0; Mismatches 586; Indels 40; Gaps 7;

QY 123 GGAAGTGAAGGGGAGAGCCATGAGGCGGCGGCTGGTGGCCGCTCCACAAC 182  
Db 180 GGGCGGGGGGGGGGGCGCTGAGATGAGGCTAGCGCGGGCTGGCCGCTCGCATAC 239  
QY 183 CGCAAGAGCTGCTGTCATCCGCGCGGAGCGGA-TCCGGGGCCGAAGCCCGCGGGA 241  
Db 240 CGGAACGAGCTGTGTGATCCGCGCGAGCTCGGAGCCCGGGCGGCGCG 299  
QY 242 GCAGAACGGGC-----AGGTGCGCAGATTGCGGCGAGCAGTGGCCTTGCCCGGC 296  
Db 300 GCGCGCGGGCGGAGCGCGCTGCGCAGATATGCGGCGAGGTCGGGCTGGCTTCAG 359  
QY 297 GGGGACCCCTTCTGTCGTCGTCAGAGTGCCTTCCCGTCTCCGGGACTGCTACGA 356  
Db 360 GGGGAGCCCTTCTGTCGTCGTCAGAGTGCCTTCCCGTCTCCGGGACTGCTACGA 419  
QY 357 TACGAGCGCGGGAGGCGACGCAACTGCCCCCAGTGAAGACTCGATACAAGCGCTC 416  
Db 420 TACGAGCGCGCGGAGGCGCTCGAGCGCTGCCAGTGAAGACCCGCTACAAGCGCTC 479  
QY 417 AAGGCTGCCCAACGTGTGACCGGTGACGAGAGGAGGACGGCGTCGATGACCTGACAC 476  
Db 480 AAGGCTGCCCGCGGGTGGTGGGACGAGAGGAGGACGGCGTCGACGACCTGAGGGC 539  
QY 477 GAGTT-----CAACTGGGACGGCCATGACTGCGAGTCTGTGGCGGAGTCC 521  
Db 540 GAGTTCGGCCTGCAGAGCGGCGCGGCCACGAGAGGACCGCGAGTACGTGCGCGAGTCC 599  
QY 522 ATGCTTACGGCCACATGAGCTACGCGCTGGAGGTGACCTTAATGCGCGCAACAGCT 581  
Db 600 ATGCTGCGGGCGCAGATGAGCTACGCGCGCGCGCA-----CGCGACCGCGAC 650  
QY 582 TTCAGCTCAACCCCAATGTTCACTCTCCACCAACGGGCAATGGTGATGACATCCA 641  
Db 651 TTCAACCCCGTCCCAACGTGCGCTCTCAACCAACGGCGCATGGTTGATGACATCCG 710  
QY 642 CCGGAGACGACGGCGCTGCTCTTTTCATG-----GGTGTGGGGGAAAGAGATA 695  
Db 711 CCGGAGACGACGGCGCTGCTGCTCTCAATGAGCGCGCGCGCGCGCAAGAGATC 770  
QY 696 CATCCCTTCTTATGCGGATCCCAAGCTTACCTGTGCAACCCAGGCTATGAGCCCATCC 755  
Db 771 CACCCGCTCCCTTCTGCTGATTCACCTTCAGTGAACCGAGATCCATGAGCCCGTCC 830  
QY 756 AAGATCTTGTCTGATATGGGTATGATGTGCTTGAAGGAACGATGAGAAATTGG 815

Db 831 AAGATCTGGCGCCTACGATATGGCGCGTGGCTGGAAGGAGAAATGAGGCGTGG 890  
QY 816 AAGCAGACAGAGAGAGATGACACGAGCGGGAATGATGGTGGTGGAT--GATGT 872  
Db 891 AAGCAGAGCAGAGCGCCTGACGATTCAGAGAGGAGGCTGGCGGTGATTTGGATGGC 950  
QY 873 GACGATGCTGATCTACCACTAATGATGTAACCAAGACACAACTGTCCAGAAAATCCA 932  
Db 951 GACGATGACGATCTGCCACTAATGATGAACTAGGACGACCATGTGTCCAGAAAATCCCT 1010  
QY 933 CTTCCATCAAGCCAGATTAACTCAATAGATGATTATCATTTATCGGCTGTGTTTG 992  
Db 1011 ATATCATCAAGCCGAATTAACTCCACAGATGATTATCGTTATCCGTTGGTGTG 1070  
QY 993 GGGTCTTCTTCCACTACCGAGTATGATCCGGTGAATGATGATTTGCTGTGGCTC 1052  
Db 1071 GGTTCCTTCCACTACCGAGTATGATCCGGCGGAAAGATGCAATTGCAATGTGGCTC 1130  
QY 1053 ATATCTGTATCTGTAATCTGTTGCCATGCTTGGATTCTGATTCATTCACCAAG 1112  
Db 1131 ATATCTGTATCTGTAATCTGTTGCCATGCTTGGATTCTGATTCATTCACCAAG 1190  
QY 1113 TGGTCCCTATGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTCCAGCAAGAA 1172  
Db 1191 TGGCTTCCAAATCGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTTGAAGAA 1250  
QY 1173 GGGCAGCCATCTCAACTTGTCTCCAATTGATTTCTTGTGATGACGTTGATCCCTTAAG 1232  
Db 1251 GGTCAACCCCTCAGCTTGTCTCAATCGACTTCTTGTGATGACGTTGATCCCAAG 1310  
QY 1233 GAACCTCCTTGGTCAACAATACTGTTCTATCTATCCTTCCGTGGAATTATCCTGTT 1292  
Db 1311 GAACCTCCTTGGTCAACAATACTGTTCTATCTATCCTTCCGTGGAATTATCCTGTT 1370  
QY 1293 GATAAGTTTCTGCTATGTTCTGATGATGCTGCAATGCTAAGCTTGAAGCATTA 1352  
Db 1371 GAGAAGTCTCCTGATATGTTCTGATGATGCTGCAATGCTTACGTTGAAGCATTTG 1430  
QY 1353 TCTGAACATCTGAATTTGCAAGAAATGGGTTCTTCTGCAACGCTACATATTTGAA 1412  
Db 1431 TCTGAACATCTGAATTTGCAAGAAATGGGTTCTTCTGCAACGCTACATATTTGAA 1490  
QY 1413 CCTGCGCTCCAGAGTGTACTTCCAAACAGAGATAGACTACTTGAAGACAAAGTGGA 1472  
Db 1491 CCTGCTCTCTGAGTGTACTTCCAAACAGAGATAGACTACTTGAAGACAAAGTGGA 1550  
QY 1473 GCAAACTTGTAGGAGAGAGAGCAATGAGAGAGAGATGAGAAATTCAGGCTGAGA 1532  
Db 1551 GCTTCATTGTAGGAGAGAGAGGCGATGAGAGAAATACGAGAAATTCAGGCTGAGA 1610  
QY 1533 ATCAATGCTTGTAGTCCAAAGCCCAAGAGTCTTGAAGAGATGCAATGCAAGAT 1592  
Db 1611 ATCAATGCTTGTAGTCCAAAGCCCAAGAGTCTTGAAGAGATGCAATGCAAGAT 1670  
QY 1593 GGAACCCCTGGCTGGAACAATGTTGATCATCTGGAATGATTCAGGTCTTCTT 1652  
Db 1671 GGAAGCCCTGGCTGGAACAATGTTGATCATCTGGAATGATTCAGGTCTTCTT 1730  
QY 1653 GGGCAAAAGCGAGGCTTGACTGTGAGGAAATGAACCTGCCAGATTTGTTATGTTCT 1712  
Db 1731 GGGCAAAAGTGGCGGTGTGATGTGAAAGAAATGAGTTGCTGCTGTTATGTCTCG 1790  
QY 1713 AGAGAGAAACGACGAGCTATAACCATATAGAAAGCTGTGCTATGAATGATGCTG 1772  
Db 1791 AGAGAGAAAGAGCCAGGTTATTAACCATACAGAAAGCTGTGCTATGAATGATGCTG 1850  
QY 1773 CGAGTCTCTGCTGTAATAAATGCTTCAATTTGTTAACTTGGATTTGATCACTAC 1832  
Db 1851 CGTGTCTGCTGCTGTAATAAATGCTGTAATCTTGAATTTGATGATCACTAC 1910  
QY 1833 ATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTATGATGAGCCCTTACTAGA 1892

Db 1911 ATCAACAATAGCAAGGCCATAAAGGCTATGTGTTTCATGATGATCCTTGTGGGG 1970  
QY 1893 AAGAAGTTGCTATGTACAGTCCCTCAAGATTGATGGATGATGCCATGACCGA 1952  
Db 1971 AAGAAAGTGTCTATGTACAGTCCCTCAGAGGTTGATGATGACAAAATGATCGA 2030  
QY 1953 TATGCTAACCGGAATGTTGCTTTTGTATCAACATGAAAGTTGATGATTCAG 2012  
Db 2031 TACGCTAACGGAACGTGTCTTTTGAACATCAACATGAAAGTTGACGATTCAA 2090  
QY 2013 GGTCCAATTATGTGTACTGTGATGTGATTTGAAGGACGATTAATGTTATGAT 2072  
Db 2091 GGACCCATTATGTGGTACTGTGATGTGTTTCAAGCGGACGACTGTATGTTATGAT 2150  
QY 2073 GCCCCCAAAACAAGAAAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGTCTT 2132  
Db 2151 GCTCCTAAACGAAGAACCAACCATCAAGACTTGCAACTGCTGGCCCAAGTGTCTC 2210  
QY 2133 TGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGACTACCAAAACCAAAACAGAGAG 2192  
Db 2211 TCTGTGCTGCTGCAAGCAAGAAATAAAGAACTACAAACCAAAACAGAGAG 2270  
QY 2193 AAAAAGTATATTTTTCAGAAAGAAAGAACCAATCCCTGCAATATGCTCTGTGAA 2252  
Db 2271 AAGAAAGATATTTTTCAGAAAGCAAGAAACCATCTCTGCAATATGCTTGTGAA 2330  
QY 2253 ATTGACGAAGCTCTCCAGAGCTGAGAAAGGCGGATTTGTAATCAACAAAA 2312  
Db 2331 ATTGATGAAGTGTCTCCAGGTGTGATATCGAAGAGCCGGAATCGTAATCAACAGAA 2390  
QY 2313 TTAGAAAGAAATTTGGCCAATCTCTGTTTGTATACATCCACACTTCTGAGAAATG 2372  
Db 2391 CTAGAGAAAGAAATTTGGCAGTCTTCTGTTTGTGCGATCAACACTTCTTGAGAACGA 2450  
QY 2373 GGAACCTTGAGAGTGCAGATCCCTGCTTCTTTTGAAGAGCTATACATGTCATTAGT 2432  
Db 2451 GGACCCCTGAGAGCGCAAGTCCAGCTTCTCTGAGAGAGCTATACATGTTATCAGC 2510  
QY 2433 TGTGTTATGAGACAAGACAGACTGGGAAAGATTGCTGATCTATGATCAGT 2492  
Db 2511 TGGGCTACGAGACAGACAGACCGACTGGGAAAGAGATTGCTGATTTACGATGATC 2570  
QY 2493 ACAGAAGATATTTCTACTGTTTCAAGATGATGTCATGTTGGCGGTCAATTTACTGC 2552  
Db 2571 ACAGAGGATATTTCTGACTGATTTAAGATGACTGCCATGCTGGCGGTCTATTTACTGC 2630  
QY 2553 ATACCTAAACGGGTTGATTCAAAGGTTCTGCACTCTGAATCTTTCAAGATCGTCTCAC 2612  
Db 2631 ATCCCGAAGCGGCTGCAATTCAAAGGTTCTGCGCTCTGAACCTTTCGACCGTCTTCA 2690  
QY 2613 CAGGTGCTTCGGTGGGCTCTTGGGCTATGAGATCTTC-TTACGAATCATTGCCCTCT 2671  
Db 2691 CAGGTCTTTCGGTGGGCTCTTGGGCTCGTGAATTTCTTTCAGNANCAATGNCACCT 2750  
QY 2672 TTGATATGGGTAATGTTGGCGGCTGAATTTTGGAAAGATTTTCTCATCACTCCAT 2731  
Db 2751 TTGATACGATACGGCGGCGGCTAAATTCCTGGAAGGTTTCTATATCACTCCAT 2810  
QY 2732 CGGTATCCTTGACATCTATTCCTCTGCTTACTGTACATGCTGCCATCTGTT 2791  
Db 2811 CGTTATCCCTGACGCTCATCTCTCTCTGCTTACTGTACCTTGCTGCCATCTGCC 2870  
QY 2792 ATTGACAGGAAATTTATCACTCCAGAGCTGAATATGTTGCCAGGCTGTGTTTCATGTC 2851  
Db 2871 GCTCAGCGGGAAGTTTATCACACCAAGGCTTACCAATGTCCGCAATCTGTTTCATGGC 2930  
QY 2852 ACTTTTATCTGATTTTCTGACGAGCATCTAGAAATGAGATGAGTGTGTTGAAT 2911  
Db 2931 ACTTTTATCTGATCTCTGTCGACCGGCGCATCTGAAATGAGTGTGAGTGTGCCAT 2990  
QY 2912 TGATGACTGTGAGGAATGAGCAGTCTGTGCTCATTTGAGGTGTGCTCTCACACTCTT 2971  
Db 2991 CGACGACTGTGAGGAACGAGCAGTCTGTGCTCATCGAGGCGTTCGCGCATCTGTT 3050

QY 2972 TGCTGNTTCCAGGACTTCTCAAGTCAATAGCTGTGTTGATACAGCTTCAACCGTAC 3031  
Db 3051 CGCGGTGTTCCAGGCGCTGCTGAAGGTGTTCGCCGGCATCGACAGAGCTTACCGTAC 3110  
QY 3032 ATCAAGGCTGAGATGATGAGAGTCTCAGAGCTATATACATTCAATGAGTACTT 3091  
Db 3111 GTGGAAGCGCGGACGACGAGAGGTTCTCGAGCTGTACAGTTCAAGTGAACACCT 3170  
QY 3092 ATTGATACCTCTTACACCTTGTCTTATGAACTTCATGTGTGTGCTGCGTTC 3151  
Db 3171 GCTGATACCCCGACACGCTCTCTGCTGAACCTTCAATCGGGTGTGCGCGGATCTC 3230  
QY 3152 AAATGCGATCAATAACGATATGAGTCAATGAGGCGCCCTCTTGGAGCTATTTGC 3211  
Db 3231 GAACGCGATCAACAACGGGTACAGTGTGAGGCGCCCTGTTCGGAGAGCTCTTCGC 3290  
QY 3212 ATTTGGGTGATGTCCATCTTTATCCCTTCTCAAGGTTGTTGGAAGCAAAACAG 3271  
Db 3291 GTTCTGGGTGATGCTCCACTGTACCCGTTCTCAAGGTTGTGAGGAGCAAGACAG 3350  
QY 3272 GACACCAACGATGTGATCGTCTGTGTCATCTGCTGCTGCTCAATCTTCTGCTCTTG 3331  
Db 3351 GACGCTGACGATGTCATCGTCTGTGTCATCTGCTGCTGCTGATCTTCTGCTCTG 3410  
QY 3332 GGTTCGATGATGATCTTCTCTTCCGGAAGATGATGTCGCTTCTTGAGAGTGTGTT 3391  
Db 3411 GGTCCGCTGACCGCTTCTCTCCCAAGAGCAGCGCCGCTCTGAGAGAGTGTGCT 3470  
QY 3392 GGAATGCAACT 3402  
Db 3471 GGACTGCAACT 3481

RESULT 4  
AY110415 3898 bp mRNA linear HTC 17-OCT-2002  
LOCUS  
DEFINITION Zea mays Cl1166\_1 mRNA sequence.  
ACCESSION AY110415  
VERSION AY110415.1 GI:21214824  
KEYWORDS  
SOURCE  
ORGANISM  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,J.L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
REFERENCE  
AUTHORS  
Coe,E.H.  
TITLE  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
COMMENT  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZMDB and may be found by BLAST  
searching at MSL, [maizemap.org](http://maizemap.org); ZMDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZMDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES  
source  
1..3898  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="Maizemap:630057"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST



assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

BASE COUNT 932 a 841 c 991 g 1054 t 80 others  
ORIGIN

Query Match 34.1%; Score 1342.8; DB 11; Length 3898;  
Best Local Similarity 64.9%; Pred. No. 1.3e-289;  
Matches 2120; Conservative 0; Mismatches 1034; Indels 111; Gaps 8;

QY 127 GTGAGGGGAGGAAAGCGATGAGAGCGGAGCGCGGCTGTCGCCGCTCCCAACCGCA 186  
DB 172 GTAGAGGAGCGCGGAGATGCGGCCAACAAAGGGATGCGGGCTCGACAAACCGCA 231  
QY 187 ACAGCTCGTCGTCATCCGCCGACGCGCGATCC-----CGGGCCGAAGCGCGCGGG 240  
DB 232 ACAGATTCGTCATGATCCGCCACGACGCGCGCGCGCGCTTAAGCCACGAAGA 291  
QY 241 AGCAGAACGGGAGGTGTCGATTTGCGGCGACGACGTCGCTTGCCCCCGCGGGG 300  
DB 292 GTGCGAATGGGACAGTCTGCCGATTTGCGGCGACTCTGCGGCGTTTACGCCATGGTG 351  
QY 301 ACCCCTTCGTGCGTCACAGATGCGCCTTCCCGCTGCGCGGACTGCTACGAATACG 360  
DB 352 ATGCTTTGTCCTGCAATGATGTCCTTCCCTGTCTGCGCGCCTTGCTATGATGATCG 411  
QY 361 AGCGCCGGGAGGCGACGACGAACTGCCCCAGTGCACGACTCGATCAAGCGCCTCAAG 420  
DB 412 AGCGCAAGGAGGGGAAACCAATGCTGCCCTCAGTGCAGACTAGATCAAGAGACGAAG 471  
QY 421 GCTGCCAAGCTGTGACCGGTGACGAGAGGAGGACGCGTCGATGACCTGACACAGAGT 480  
DB 472 GTAGCCTTCGAGTTCATGATGAGATGAGAAAGATGATGACCTAGACAAATGAT 531  
QY 481 TCAACTGGAGCGGCCATGACTCGAGTCTGTGCCGAGTCCATGCTTACGCGCACATGA 540  
DB 532 TCAACTACAA-----GCAAGGAGTGGGAAGGCGCCCAAGT 567  
QY 541 GCTACGCGCGGTGAGGTGACCCCTAATGCGCGCCACACAGCTTTCCAGCTCAACCCCAAT- 599  
DB 568 GGCAGCTGCAAGAGATGACCGTGAATGCTCTTCACTGTCTGCCCATGACCCACACCATC 627  
QY 600 --GTTCCACTCTCACCAACGGGCAAAATGCTGATGACATCCCAACGGAGCAGCAGCGC 657  
DB 628 GGATTCACGCGCTGACAAAGCGGTCAACAGATATCTGAGAGATCCCTGATGATCCCTTG 687  
QY 658 TGGTGCCTTCTTTCATGGTGGGGGAAAGAGATACATCCCTTCTTATGCGATC 717  
DB 688 ACCGTATTTCTATC-----CGCAGTCAACATCGAGCTATGTTGATC 729  
QY 718 CCAGCTTACTGTGCAACCCAGGTCTATGAGCCCATCCAGAGATCTTGTGCTGATAGGT 777  
DB 730 CAAGCGTCCAGTCTCTGTAGAGATGTGGAACCCCTGAAAGACTTGAATTCCTATGGGC 789  
QY 778 ATGATGCTGTGCTTGAAGGAACGATGAGAAATGGAAGCAGACAGACAAGAGAGATGC 837  
DB 790 TTAATAGTGTGACTGGAAGGAAGAGCTGAGAGCTGAGAGGTTAAACAGACNNNNNTA 849  
QY 838 ACCAGACGGGAATGATGTTGTTGATGATGTTG----- 873  
DB 850 TGTTCGAGTGAATAATTAATATCCAGAGCTAGAGGAGAGACATGAGGGGACTGGCT 909  
QY 874 ---ACGATGCTGATCTACCACTAATGATGAGCAAGACAACAATGTCAGGAAATTC 930  
DB 910 CAAATGAGAAAGATATGCAATGTTGATGATGACGCTACCTTTGAGCCGATCGTGC 969  
QY 931 CACTTCATCAAGCCAGATTAATTCATATAGAGATGATTAATTCAGCTTGTGTTT 990  
DB 970 CAATTTCTCAAAACAGCTCAACCTTTTACCGGGTAGTATCATCTCCGTTTATCATCC 1029  
QY 991 TGGGTTCTTCTTCACTACCGAGTATGATCCGGTGAATGATTTGCTTGTGCG 1050

DB 1030 TGTGCTTCTTCCAGTATCGTATCGATCCAGTGCCTGATGCTTATGATGTGGC 1089  
QY 1051 TCATATCTGTATCTGTGAATCTGGTTTGCCATGTCTTGATCTTGTATCAATTCCAA 1110  
DB 1090 TAGTATCTGTATCTGCGAGGCTGTGTTGCTTGTCTGCTTGTATGATCAAGTTCCAA 1149  
QY 1111 AGTGTTCCTTATGAGAGAGACTTAAGACCGGCTGTCACTGAGGTTCGACAAGG 1170  
DB 1150 AATGTATCAATCAACCGTGAAGACATATCTGACAGGCTTGATGAGATGATAGAG 1209  
QY 1171 AAGCCAGCATCTCAACTTGTCTCAATGATTTCTTGTACATGAGTTGATCCCTTAA 1230  
DB 1210 AGGAGAGCCATCACAGCTGCTCCCATGATGCTTGTGATGATGATGATGATGATG 1269  
QY 1231 AGAACCTCCTTGTGTCACACAATACTGTCTATCTATCTATCTTTCGTTGATGATTCCTG 1290  
DB 1270 AGAACCTCCTGATCACAGCCACACTGTTTGTCCATCTTGTGATGATGATGATGATG 1329  
QY 1291 TTGATAAGTTCTTGTCTATGTTCTGATGATGCTGCTCAATGCTAACGTTGAAGCAT 1350  
DB 1330 TTGACAAAGTGTATGCTATGTTCTGATGATGCTCAGCTATGCTGATCTTTGAGTCTC 1389  
QY 1351 TATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTCTGCAACGCTACATATTTG 1410  
DB 1390 TCTCTGAACCGCGCAATTTGCTAGAAAGTGGTTCCTTTGTGAAGACCAATATTTG 1449  
QY 1411 AACCTGCGCTCAGAGTGTACTTCCACAGAAATAGACTTACTTGAAGACAAGGTGG 1470  
DB 1450 AACCAAGAGCTCCGAATTTTACTTGTCTNNNNNTAGATTACCTGAAGACAATAATTC 1509  
QY 1471 CAGCAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATGAGAAATCAAGGTGA 1530  
DB 1510 AACCTTATTTGTTAAGAAAGACGCGCAATGAAGAGGATGATGAAGAAATTCAAAGTAA 1569  
QY 1531 GAATCAATGCTTGTGTCACCAAGCCCAAGAAATTCCTGAAGAGAGATGACAAATGCAAG 1590  
DB 1570 GAATCAATGCTTGTGTCACCAAGCAAGAAATGCTTGAAGAGGAGGATGACCAATGCTG 1629  
QY 1591 ATGGAACCCCTGCGCTGAAACAATGCTGATCATCTGGAATGATTCAGGCTTCC 1650  
DB 1630 ATGGAACCTGCTGCGCTGGAATTAACCTAGGAGCCATCTGCGATGATTCAGGCTTCT 1689  
QY 1651 TTGCCAAAGCGAGCGCTGACTGTGAGGAAATGAATGACCGACGATGCTTATGTTT 1710  
DB 1690 TGGGACAGTGTGGGCTCGACACTGATGAAATGATTACACGCTTGTCTATGCTCT 1749  
QY 1711 CTAGAGAGAAACGACGAGCTATTAACCATCAATAAGAAAGCTGTGCTATGATGATGCT 1770  
DB 1750 CTCGTGAAGAAAGACGAGCTTTCAGCATCAACAAGAGCTGTGCAATGATGCGCTGA 1809  
QY 1771 TCCGAGTCTCTGCTGTAATAACAATGCTCAATTTGTTAACTTGAATGATGATCACT 1830  
DB 1810 TTCGTGATCTGCTGCTGACAAATGCTGCTATCTTCAATGATGATGATGATGATGATG 1869  
QY 1831 ACATCAACACAGCAAGCTATTAAGGAAGCAATGTTTATGATGATGATGATGATGATGATG 1890  
DB 1870 ACTTCAATAGCAGCAAGCTCTTGAAGAGCAATGCTTCAATGATGATGATGATGATGATG 1929  
QY 1891 GAAAGAAAGTTGCTATGTAAGTCTTCCCAAGATTTGATGAGATGATGATGATGATGATG 1950  
DB 1930 GAAAGAAACTTGTATGTAACAATTTCCACAGAGATTTGATGAGATGATGATGATGATG 1989  
QY 1951 GATATGCTAACCGGAATGTTGCTTTTGTGATTAACAATGAAAGGTTGATGATGATGATG 2010  
DB 1990 GATATGCTAATCGGAACATAGTCTTCTTGTATTAACAATGAAAGGCTGATGATGATGATG 2049  
QY 2011 AGGTTCATTTATGTTGCTACTGATGTTGATTTAGAAAGCAGGATTAATGATGATGATG 2070  
DB 2050 AGGTTCAGTCTACGTGGAACAGAGATGCTGTTCAATAGGCAAGGCTTTGATGATGATG 2109  
QY 2071 ATGCCCCCAAAACAAAGAACCAACCATCAAGACTTGCACACTGCTGCGCAAGTGTGCT 2130

D	b		2110	ATCC-----TGTTGACTGAAGCTGATCTGGACCT	2141
O	y		2131	TTTGTGTTGCTGCTTTGGCAATAGAAGCAAAAAGAAAGTAACAACCACCAAGCAGAGA	2190
D	b		2142	AACATTGTTGTTAAGAGCTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATAAGAT	2201
O	y		2191	AGAAAAAGTTATATTTTCAGAAGAAGAGAACAATCCCCCGCATATGCTCTTGGTG	2250
D	b		2202	AGTCAAAGCCGTA--TTATGAAGAGAACAAATCTTCAGCTCCCACATCTTCAACATGGAAAG	2259
O	y		2251	AAATTGACGAAGCTGCTCCAAGAGCTGAGATGAAAAGCCGGTATTTGTAATCAACAA	2310
D	b		2260	ACATCGAGAGGATATTGAAGSTTACGAGATGAAAGGTCAGTCTTATGTCCAGAGGA	2319
O	y		2311	AATTAGAAAAAATTTGCCAATCTTCTGTTTTTTGTATCATCCACACTTCTCGAATG	2370
D	b		2320	AATTGAGAAACGCTTGTGTCAGTCTCCAATCTTCATGTCATCCACCTTATGATCAAG	2379
O	y		2371	GTGGAACCTTGAAGAGTGCAGTCCCTGCTTCTCTTTGAAGAAGCTATACATGTCATTA	2430
D	b		2380	GTEGCATACCACCTTCAACAAACCAGCTTCTCTACTAAGGAAGCTATTCATGTCATCA	2439
O	y		2431	GTTGTGTTATGAAGACAGACAGACTGGGAAAAAGATTGGCTGATCTATGATCAG	2490
D	b		2440	GTTGTGATATGAGACAAACTGAATGGGAAAAAGATTGGCTGATCTATGTTGAG	2499
O	y		2491	TTACAGAAGATATTCTAATGTTTCAAGATGATGTCATGTTGGCGTCAATTTACT	2550
D	b		2500	TTACGAGGATATTCTGACTGGGTTAAANNNNNNNNNAGGGGCTGGCAATCAATCTACT	2559
O	y		2551	GCATACCTAAACGGGTTGCATTTCAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTC	2610
D	b		2560	GCATGCCACACGACCTTGTTTCAAGGGTTCTGCACCAATCAATCTTCTGATGTCCTTA	2619
O	y		2611	ACCAGTGCTTCCGTGGGCTCTGGGCTATTTGAGATCTTCTCAGCAATCATTGCCCTC	2670
D	b		2620	ATCAGGTGCTCCGTGGGCTCTGGGCTAGTGAATTTCTGCTTAGCAGACATGTCTTA	2679
O	y		2671	TTTGTATGGTATGTTGGCGGCTCGAAATTTTGGAAAGATTTCTCATCACTCACTCCA	2730
D	b		2680	TCTGSTATGTTACAATGACGATTTGAAGCTTTTGGAGAGCTGGCTTACATCAACACCA	2739
O	y		2731	TTCGTATCTTGGACATCTATTCCCTCTTGGCTTACTGTACATTGCCCTGCCATCTGTT	2790
D	b		2740	TGTGTTATCCAATCACATCTGTTCCGCTTATCGCCTATGTGTGCTTCTCTGTAATCTGTC	2799
O	y		2791	TATTGACAGGAAATTTATCACTCCAGAGCTGAATTAATGTGCCAGCCGTGTGTTCAATGT	2850
D	b		2800	TTCTTACCAATAAATTATTCATCTCTGAGATTAGCAATATATGCTGGATGTTCTTCATTC	2859
O	y		2851	CACTTTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGAGTGGTGTGAA	2910
D	b		2860	TTCTTTTCCCTCCANNNNNNGCCACTGATATTGGACCTTAGATGAGTGGTGTGGCA	2919
O	y		2911	TTGATGACTGTGAGAGAAATGACAGTCTCGGTCATTGGAAGTGTGTCTCACACTCT	2970
D	b		2920	TTGAAGATTGGTGAAGAAATGAGCAGTTTGGGTTATGTGTGGCACCTCTGCCCATCTCT	2979
O	y		2971	TTGCTGTGTTCCAGGACTTCTCAAGTCAATAGCTGGTGTGATTACAAGCTTCAACCGTGA	3030
D	b		2980	TTCAGTGTTCAGGGTCTGCTGAAGTGTGGCTGGCATGTATTACCAACTTCACAGTTA	3039
O	y		3031	CATCAAAGG---GTGAGATGATGAGAGTCTCAGAGCTATATACATTCAAAATGACTA	3087
D	b		3040	CCTCAAAGGCATCTGATGAGATGGCGCACTTGTGCTAGCTATATGTGTTCAAGTGAACA	3099
O	y		3088	CCTTATTGATACCTCCTACCACTTGCTTCTATTGAAC TTCAATTGCTGTGTCGCTGGCG	3147
D	b		3100	GTTTGCTCATTCCTCCGACCACTGTTCTGTTCATTAACTGTGCGAATGTGTGACAGAA	3159
O	y		3148	TTTCAAATGCGATCAATTAAGGATATGAGTCATGGGCCCCCTCTTTGGGAAGCTAATCT	3207
D	b		3160	TTTCGTATGCCATTAACTGAGCTACCAATCCTGGGGTCCGCTCTTTGGAAAGCTGTCT	3219

QY	3208	TTGCATTTTGGGTGATGTGCCATCTTAATCCCTTTCTCAAGGTTTGTTGAAGGCATAA	3267
Dp	3220	TCTCATCTGGGTGNTTCCTCCATCTGTACAGCTTCCTCAAGGTCATGGAAGGCAGA	3279
QY	3268	ACAGGACACCACGATTGTCATCGTCTGGTCCATTCTGCTGCTTCATCTTCGCTCC	3327
Dp	3280	ACCGCACACCAACAATCGTCATTGTCTGGTCCATCCTCTTGATCTATCTTCCTTGC	3339
QY	3328	TTTGGGTTCCGATGATCCTTTTCT	3352
Dp	3340	TGTGGGTGAAGATCGATCCTTTTCAT	3364

RESULT 5			
AY108113			
LOCUS	AY108113	3763 bp	mRNA
DEFINITION	Zea mays PCO126465 mRNA sequence.		linear
ACCESSION	AY108113		
VERSION	AY108113.1	GI:21211191	
KEYWORDS	HTC.		
SOURCE	Zea mays		

REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
AUTHORS	1 (bases 1 to 3763) Halney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL	Unpublished (2002)
REFERENCE	2 (bases 1 to 3763)
AUTHORS	Coe, E.H.

**JOURNAL** Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

**COMMENT** If you are interested in getting corresponding physical clones

**COMMENT**  
If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR, [www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES	Location/Qualifiers
Source	1. .3763

```

source
1. .3763
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637367"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/Dupont Cornsensus
library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for
overgo addressing of BACs in conjunction with the M
Mapping Project"

```

BASE COUNT	863 a	984 c	1011 g	904 t	1 others
ORIGIN					

Query Match	32.8%;	Score 1289.4;	DB 11;	Length 3763;
Best Local Similarity	64.7%;	Pred. No. 1.2e-277;		
Matches 2112; Conservative	1;	Mismatches 1042;	Indels 108;	Gaps 9;

Qy	210	GACGGCGATCCCGGGCCGAA	CGCGCGGGAGCAGAA	CGGCGAGTGTCCAGATTGC	269
Db	344	GCGCGCGGACGGCGGTGA	AGTGGGGAGGCGCGGTG	CGGACAGGTGTGCCAGATTCTGC	403

**OY** 270 GCGACGACGTGGCCTTGCCTTCGCCCGGAGCAACCCTTCGTGGCGTGCAAAGATGCGCC 329  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
**Db** 404 GCGCACGGCGTGGGCAACACGCGGAGAGGGGACACTTTCGCCCGCTTGCGAAGTCTGCGGG 463  
||||| ||||| ||||| ||||| ||||| ||||| |||||

OY		330	TTCCCGCTGCGGGACTGTCAAGAAATACGAAGCCCGGAGGGCAGCAGAACTGCCCC	389
Db		464	TTTTCCGATGTGCCCCCTTGCTACGAGTACGAGCGCAAAGACGGCACGCGCTGCCCC	523
OY		390	CAGTGAAGACTGCATTACAAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGAG	449
Db		524	CAGTGAAGACCAGAATAACAAGCGCCACAAAGGGAGCCCCGGCATCCCTGGGAGGAAGA	583
OY		450	GAGGACGGCGTCGATGACCTGGACAACGAGTTCAACTGGGACGGCCATG-----ACTCG	503
Db		584	GACGACACTGATGCCGATAGCGAATTCAATTACTCTTGCAATCTGGCAATGAGGACCAAGA	643
OY		504	CAGTCTGTGCGCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGAGGTGACCT	563
Db		644	CAGAAGATTGCCGACAAATGCGCAGCTGCGCATGAACTTTGGGGCAGCGGGATGTT	703
OY		564	AATGGCGCGCAACAAGCTTTCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAA	623
Db		704	GGTCGCCCCAAGTATGACAGTGGCGAGATCGGGCTTAACAAGTATGACAGTGGCAGATT	763
OY		624	ATGTGATGATCATCCCACCGGAGCAGCAGCGCGTGTGCTTCTTCATGGGTGTGG	683
Db		764	CCTCGGGATACATCCCATCAGTCACTAACAGCCAGATCTCAGAGAAATCCTGTGTCT	823
OY		684	GGAAGAGGATACATCCCT--TCCTATGCGATCCAGCTTACCCTGTGCAACCCAGG	740
Db		824	TCCCTGACCATCAATGATGTGCCCAACTGGAACATTGGCAAGCGTGTCCATTTC	883
OY		741	TCTATGACCCCATCCAAGATCTTGCTGATATGGTATGTTAG-----TGTGCT	791
Db		884	TATGTGAACCATTCGCCCAATCCGTCAAGGAGTCTCTGTGATGGAAATGTTGCC	943
OY		792	TGGAAGGAACGATGAGAA TTGGAAGCAGAGACA GA-----	829
Db		944	TGAAAAGAGAGGTTGATGCTGAAAAATGAAGCAGGACAAGGGGAGATTCCCATGACG	1003
OY		830	-----GAGATGCAACGAGCGGGAA TGATGTTGTTGATGATGTTG-----	873
Db		1004	AATGCACAAGCATGTCTCTCTGAGGTCGGGGTGTGGTATATTGATGATCAACT	1063
OY		874	-----ACGATGCTGATCTACCACTAATGATGAAGCAACAACCTGTCCAGAAA	926
Db		1064	GATTACAACATGGAAGATGCTTATTGAACGAGAACTCGACAGCCTCTATCTAGAAA	1123
OY		927	ATTCCACTTCCATCAAGCCAGATTAA TCCATATAGATGATTA TCATTATTCGGCTGTG	986
Db		1124	GTTCCACTTCTTCTCCAGAGATAAATCCATACAGATGTGTCA TTGTCTCGATTGATT	1183
OY		987	GTTTGGGGTCTTCTTCCACTACCGAGTGATGATCCGCTGAATGATGATTCCTTTG	1046
Db		1184	GTTCTAAGCATCTTCTTGCACTACCGTATCAMAATCTGTGCGCAATGATACCAATTA	1243
OY		1047	TGGCTCATATCTGTTATCTGGAATCTGTTGGCATGTCTTGATTTCTTGATCAATTC	1106
Db		1244	TGGCTTCTATCTGTTATATGTAGATCTGTTTGTCTTTCGTGATATTGATGATGTTT	1303
OY		1107	CCAAAGTGTTCCTATTGAGAGAGAGACTTACCTTAGACCGGCTGTCACTGAGTTGAC	1166
Db		1304	CCTAAGTGTTCCTCAATCAACGGGAGACGTACTTGATAGGCTGGCATTAAGTATGAC	1363
OY		1167	AAGGAAGCCAGCCATCTCAACTGTCTCCAATTGATTTCTTGTCAGTACGGTTGATCCC	1226
Db		1364	CGGGAAGGTGAGCCATCTCA GTTGCTGTGTCATTTTCGCAGTACAGTGCACCA	1423
OY		1227	TTAAAGGAACCTCCTTGGTCAACA AATACTGTTCTPATCTATCCTTTCGTGATTAAT	1286
Db		1424	ATGAAGAGACCTCCTCTTGTCACTGCCAATAACGCTGTCAATCTTGCTGTGATTAAC	1483
OY		1287	CCTGTTGATTAAGGTTCTTGTCTATGTTTCTGATGATGTTGCTGCAATGCTAACGTTGAA	1346
Db		1484	CCTGTGATTAAGGTTCTTGTCTATGTATCTGATGATGAGCTGCCGATGCTGACATTGAT	1543
OY		1347	GCAATTAICTGAACAATCTGAATTTGCAAGAATAAGGTTCTTTCTGCAACGGTACAAT	1406

Db	1544	GCAC	TAGCTGAGACTTCAGAGTTTGCTAGAAAATGGGTACCATTTGTTAAGAGTACAAC	1603
Qy	1407	ATTGAA	CCCTGCGCTCCAGAGTGTACTTCCACAGAAAGATAGACTACTTGAAAGACAAG	1466
Db	1604	ATTGAA	CCCTGAGCTCTGAAATGGTACTTCTCCAGAAAATTTGATTACTTGAAGACAAA	1663
Qy	1467	GTGGCAG	CAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAGATATGAGAAATTCAAAG	1526
Db	1664	GTGCAC	CCCTTCATTGTGTTAAAGACCGCGGGCCATGAAGAGAAATATGAAGAAATTCAAA	1723
Qy	1527	GTGAGA	ATCAATGCGCTTAGTTGCCAAAGCCAGAAAAGTTCTCTGAAGAAAGATGCAATG	1586
Db	1724	GTTAGG	TAATGCGCTTGTGTTGCTAAGGCACAGAAAGTTCTCTGAGAAAGATGATCATG	1783
Qy	1587	CAAGAT	GGAACCCCTCGCTGGAACAATGTTGCTGATCATCTCGAATGATTCAGCTC	1646
Db	1784	CAAGAT	GGCACACACCATGGCCAGGAAACAATACCMGGGACCATCTGGAATGATTCAGGTT	1843
Qy	1647	TTCTTG	GCCAAAGCGGCTTGACTGTGAGGGAATGAACCTGCCACGATTGGTTAT	1706
Db	1844	TTCTTG	GTCACAGTGGTGGCCTTGATACTGAGGGCAATGAGCTACCCGTTTGGTCTAT	1903
Qy	1707	GTTTCT	AGAGAAAAGACGACGAGCTATTAACCATCAATAAGAAAGCTGTGCTATGAATGCA	1766
Db	1904	GTTTCT	CGTGAAGGCTCCTGGATTCCAGCATCAAGAAAGCTGTGCCATGATGCT	1963
Qy	1767	TTGGT	CCGAGTCTGTGCTACTAACAAATGCTCCATATTTGTTAACTTGATTTGAT	1826
Db	1964	CTTGTT	CGTGTCTCAGCTGTGCTTACCAATGAGCAATACATGTTGAATCTTGATTGTGAT	2023
Qy	1827	CAC	TACATCAACAACAAGGCTATTAAGAGCAATGTTTATATGATGACCCCTTTA	1886
Db	2024	CAC	TACATTAAACAACGTAAGGCTCTCAGGGAAGCTATGTGCTTCTATGAGCCCTAAC	2083
Qy	1887	CTAGA	AAAAGAGTTGCTATGTACAGTTCCCTCAAGATTGTGATGGATTGATCGCAT	1946
Db	2084	CTAGA	AGAGAGTGTCTGTACGTCCAGTTTCCCAAGATTTCGATGGCATTGACAGAAAT	2143
Qy	1947	GACCGA	TATGCTAACCCGGAATGTTGCTTTTGTGATATCAACATGAAAAGTTTGGATGGT	2006
Db	2144	GATCGA	TATGCCAACAGAACACCGGTTTTTCGATATTTAACTTGAGAGGCTTGTGATGGC	2203
Qy	2007	ATTCAG	GGTCCAATTATGTGGTACTGATGTGATTTGAAGGACGACATTATATGGT	2066
Db	2204	ATCCA	AGGACCAATTATGTCGAACTGGCTGTGTTTCAACCGAACAGCTCTATATGGT	2263
Qy	2067	TATGAT	CCCCCAAAACAAGAACCCACCATCAAGACTTGCAACTGTGCCCCAAGTGG	2126
Db	2264	TATGAG	CCCCCAATTAAAGCAGAAAG-----AAGGGT 2293	
Qy	2127	TGCTTT	TGCTGTGCTGTGCTTTGGCATAGGAAGCAAAAGAACTAACCAACCCAAAACA	2186
Db	2294	GGTTT	CTTGTCACTATGTGGCGGTAGG-----AAGAAAGCAAGCAAAATCAAAAGAG	2347
Qy	2187	GAGAA	AAAAAGTTATATTTTCAAGAAAAGAGAAACAATCCCTGCATATGCTCTT	2246
Db	2348	GGCT	CGGACAAGAAAGTCGAGAAGCATGTGACAGTTCTGTGCCAGTATTTCAACCTT	2407
Qy	2247	GGTGA	AATTGACGAAGCTGCTCCAGAGC-----TGAGAAATGAAAAAGCCGATATTGTA	2300
Db	2408	GAA	GATATATAGAGAGGAGATTGAAGCGCTGATTTGACGACGAGAAATCACTTCTTATG	2467
Qy	2301	AATCA	CAAAAAATTAGAAAAGAAATTTGGCCAACTCTTCTGTTTTTGTATACATCCACACTT	2360
Db	2468	TCTCA	AATGAGCCTGGAGAAAGATTTGGCCAGTCCGCAAGGTTTGTGCTCCACTCTG	2527
Qy	2361	CTCGA	GAATGTGGAACTTGAAGAGTGCAGAGTCCCTGCTTCTCTTTGAAAGAGCTATTA	2420
Db	2528	ATGGA	GTATGTGTGTTCTCAGTCCGCAACTCCGAGTCTCTTCTGAAAAGAGCTATTC	2587
Qy	2421	CATG	TCATTAGTTGTGTTATGAAGACAGACAGACTGGGGAAGAAAGATTGGCTGATC	2480



Db 2588 CATGTATAAGCTGTGCTATGAGACAGACTGAATGGGAACGTGATCCGGTGATC 2647  
Qy 2481 TATGATCAGTTACAGAGATATTTAACTGGTTCAAGATGCATTGTGATGTTGGCGG 2540  
Db 2648 TACGGTTCTGTCAGAGAACATTTCTACCGGATTCAGATGCACGGCGGAGGCTGGCGG 2707  
Qy 2541 TCAATTACTGCATACCTAAACGGGTTGCATCAAGGTTCTGCACCTCTGAATCTTTCA 2600  
Db 2708 TCGATCTACTGCATGCCCAAGCGGCCAGCTTTCAAGGGGTCTGCCCATCAATCTTTCC 2767  
Qy 2601 GATCGTCTTCAACGAGTGTCTCGTGGGCTCTGGTCTATGAGATCTTCTTCAAGCAAT 2660  
Db 2768 GACCGTCTGAACAGGTGCTCCGGTGGCTCTGGGTCCGTGAGATCTCTTCAAGCCGG 2827  
Qy 2661 CATGCCCTCTTGGTATGGTATGGTGGCGGTCTGAATTTTGAAGAATTTTCTCTAC 2720  
Db 2828 CACTGCCCTGTGTGTCAGGCTACGAGGGCGGCTCAAGTTCCTGAGAGATTCGCGTAC 2887  
Qy 2721 ATCAACTCCATCGTGTATCTTGGACATCTATCCCTCTTGGCTTACTGTACATTCCT 2780  
Db 2888 ATCAACACCAACCATCTACCCGCTCACGTCATCCCGCTTCTCATCTACTGCATCTGCC 2947  
Qy 2781 GCCATCTGTTTATTGACAGGGAATTTATCACTCCAGAGCTGAATATGTTGCCAGCCTG 2840  
Db 2948 GCCATCTGTTCTGTCTACCGGAAGTTCAATTCAGAGATCAAGCACTTCGCCAGCATC 3007  
Qy 2841 TGGTTCATGTCACCTTTTATCTGCACTTTTGTCTACGAGCATCTAGAAATGAGATGAGT 2900  
Db 3008 TGGTTCATCTCCCTCTTCACTCGATCTTCCGCAACGGGATCTTGAGATGAGTGGAGC 3067  
Qy 2901 GGTGTGAATGTATGACTGTGAGGAATGACAGTCTGGGTCAATTGAGGTGTGTCC 2960  
Db 3068 GGGGTGGGATCGACGAGTGTGAGGAACGAGCAGTCTGGGTGATCGGGGGCATCTCC 3127  
Qy 2961 TCACACCTCTTGTCTGTGTCCAGGACTTCTCAAGTCAATAGCTGTGTTGATACAAGC 3020  
Db 3128 GCGCACCCTTTCGCCGTGTTCAGGGGCTGTCTCAAGTGTGCGCGGATCGACACCAC 3187  
Qy 3021 TTCACCGTGACATCAAAAG--GTGAGATGATGAGAGTTCTCAGAGCTATATACATTC 3077  
Db 3188 TTCACCGTGACCTCCAAAGCCTCGAGAGAGAGCGGCACTTCGCGAGCTGTACATGTTCC 3247  
Qy 3078 AAATGACTACCTTATTGATACCTCTCAACCTTGCTTCTATGAATTCATTGATGTG 3137  
Db 3248 AAGTGAAGACGCTCCTGATCCGCGCCACCAACCATCTGATCATCAACTGTGCGGCTC 3307  
Qy 3138 GTGCTGGCGTTCAATATGCGATCAATAACGATATGATCATGCGGCCCCCTTTGGG 3197  
Db 3308 GTCGCGGCACTCTCTACGCCATCAACAGCGGATACCAAGTGTGCGGCGGCTCTTGGC 3367  
Qy 3198 AAGCTATCTTTCATTTGGGTGATGTCATCTTATCCCTTTCTCAAAAGTTGGTT 3257  
Db 3368 AAGCTCTTCTTGGCCTTCTGGGTCACTGTCACCTGTACCCGCTCTCAAGGGCCTCATG 3427  
Qy 3258 GGAAGCAAAACAGACACCAACGATTCATGTCGTGTCATCTGCTGCTTCAATC 3317  
Db 3428 GGCAGGCAAAACGACACCCGACCAATGTCGTGTCGTGAGGCAATCTGCTGCGTCCATC 3487  
Qy 3318 TTCTCGCTCTTGGGTTGCGATGATCTTCTCTTGCAAGATGATGTCGCTTCTT 3377  
Db 3488 TTCTCTCTGCTGTGGTTCGATGACCCCTTCAACACCCGCGTCACTGGCCCGATACC 3547  
Qy 3378 GAGAGTGTGCTTGGATTGCA 3400  
Db 3548 CAGACGTGTGGCATCAACTGCTA 3570

RESULT 6  
LOCUS AY110079 3696 bp mRNA linear HTC 17-OCT-2002  
DEFINITION Zea mays CL1164\_1 mRNA sequence.  
ACCESSION AY110079  
VERSION AY110079.1 GI:21214162

KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 3696)  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 3696)  
AUTHORS Coe,B.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.  
FEATURES  
Source  
1..3696  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="maizedb:630055"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Cornsensus library"  
/note="this sequence is part of a project of BSI assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 831 a 766 c 880 g 945 t 274 others  
ORIGIN

Query Match 30.8%; Score 1212.4; DB 11; Length 3696;  
Best Local Similarity 61.2%; Pred. No. 2e-260;  
Matches 2004; Conservative 0; Mismatches 1162; Indels 108; Gaps 8;

Qy 210 GACGGCATCCGGGCGAAGCCGCGGAGACGAGACGAGGAGTGTGCCAGATTGC 269  
Db 257 GACGGCGGCGACGCGCAAGATTTCGGGAGCATGTGCGGGCAGGTGTGCCAGATCTGC 316  
Qy 270 GCGGACGAGCTGCGCTTGGCCCCCGGGGAGACCCCTTCTGCGGTGCAACGAGTGC 329  
Db 317 GCGGACGCGGTGGGCAACCGCGCGGAGCGGACCTCTTCAACCGCTGCGAGCTGCGGC 376  
Qy 330 TTCCCCGTCTGCGGAGCTGCTAAGATACGAGCGCGGAGGAGGAGGAGGAGGAGGAG 389  
Db 377 TTCCCCGTGTGCGGCGGAGCTGCTAAGATACGAGCGGAGGAGGAGGAGGAGGAGGAG 436  
Qy 390 CAGTGAAGACTGATACAGCGGCTCAAGGGCTGCGCAAGCTGTGACCGGTGACGAGAG 449  
Db 437 CAGTGAAGACTAAGTACAGCGGCGCACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 496  
Qy 450 GAGAGCGGCGTGCATGACCTGACACAGATTCACT-----GGAGCGGCAT 497  
Db 497 GAGGATGTGATGTGACGATGTGAGTACTACCACTAACCAAGCATCTGCAACAGAGAT 556  
Qy 498 GACTCGAGTGTGTGGCGGAGTCCATGCTTACGGCCACATGAGCT----- 543  
Db 557 CAGAAACAAAGATTGCTGAGAGATGCTCACTTGGCGGACAACTCACTGCGAGTGAT 616  
Qy 544 -----ACGGCGGTGAGGTGACCTTAATGGCGGCGCAAGCTTTC 584  
Db 617 ATTGGCCTGCTAAGTATGACAGCGGTGAATTTGGGCGATGGGAAGTATGACAGTGTGAG 676

QY 585 CAGCTCAACCCCAATGTTCCACTCCTCACCACCGGCAATGSGTGAATGACATCCACCG 644  
| | | | |  
Db 677 ATCCCTCGTGATATATCCCGTCACTAATCATAGCCAGATCTCAGAGAGATTCCTGGA 736  
QY 645 GAGCAGACACGCGCTGCTGCTCTTTCATGGGTGTTGGGGAAGAAGATACATCCCTT 704  
| | | | |  
Db 737 GCTTCCCTGATCATATGATGTCCTCTGTTGGACATGGCAGCGGTGACATCATTTT 796  
QY 705 CCTTATGCGGAT-----CCCAAGTTACCTGTGCAACCCAGGTCTATGACCCTAT 753  
| | | | |  
Db 797 CCTTATGTAATCATTTCTCCAAACCCATCGAGGGAGTTCTCCGGTAGCCTTGCCATGTT 856  
QY 754 CCAAGGATCTTGCTGCATATGGGTATGTAAGTGTGCTTGGAAGGAACGGATGAGAAAT 813  
| | | | |  
Db 857 GCATGAAAAGAGAGGGTGAATGGAATAATGAAGATAAAGTGCAATTCTTATGACC 916  
QY 814 GGAAGCAGAGACAAGAGAGATGCAACGAGGGA-----ATGATGGTGGTGTGAT 866  
| | | | |  
Db 917 AATGGAACAAGCATTGTCTCCATCAGAAAGGCGTGAGTGTGATATTGATGCTTCTACT 976  
QY 867 GATGGTACGATGCTGATCTACCACTAATGATGAAGCAAGCAACAACCTGTCCAGAAA 926  
| | | | |  
Db 977 GATTAATAACATGGAAGATGCCCTTACTGAATGATGAAACTCGGCAACCTCTATCTAGAAA 1036  
QY 927 ATTCACCTTCCATCAAGCCAGATTAATCCATATAGATGATTAATCATATTCCGCTGTG 986  
| | | | |  
Db 1037 GTGCCAATTCCTTCATCCAGAAATAATCCGTACAGAAATGTCATTTGTCTACGTTTGGCT 1096  
QY 987 GTTTGGGGTCTTCTTCCACTACCGAGTGAATGATCATCCGGTGAATGATGATTTGCTTTG 1046  
| | | | |  
Db 1097 GTTCTATGCAATATCTTGCGCTACCGTATCAACATCTGTGAACAATGCAATATCCACTG 1156  
QY 1047 TGGCTCATATCTGTATCTGTGAAATCTGTTTGCCATGCTTGGAATCTTGATCAATTC 1106  
| | | | |  
Db 1157 TGGCTTTATCCGTATATGAGATCTGGTTTGTCTTGCTGGAATTTTGATCAGTTTC 1216  
QY 1107 CCAAGTGTCTCCCTATTGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGGTGAC 1166  
| | | | |  
Db 1217 CCAAGTGGTCCCCAATCAACCGTGAACATACCTGTATAGACTGGCTTTAAGGTATGAC 1276  
QY 1167 AAGAGGCCAGCCATCTCAACTGTCTCAATTGATTTCTTGTCAAGTACGTTGATCCC 1226  
| | | | |  
Db 1277 CGAAGGTGAACCATCTCAATTAGCTCCTGTGATANNNNNGTCAGTACTGTGATCCA 1336  
QY 1227 TTTAAGGAACCTCTTGTGTGTCACACAATACTGTCTATCTATCTTCCGTTGATAT 1286  
| | | | |  
Db 1337 ATGAAGAGCCTCTCTGTGTCACTGCAATACTGTGCTTCCATCTTGTCTGTGATAT 1396  
QY 1287 CCTGTGATAAGGTTTCTGTATGTTTCTGATGATGTTGCTGCAATGCTAACGTTGAA 1346  
| | | | |  
Db 1397 CCGGTGACAAGGTATCTGTATGTTTGGATGATGAGCTGTATGCTGACTTTTGAT 1456  
QY 1347 GCATTAATGAAAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACCGGTACAT 1406  
| | | | |  
Db 1457 GCTCTCTGAAAATTAGAGTTGTAAGAAATGGGTTCCGTTCTGTAAAGATACAAAC 1516  
QY 1407 ATTGAACCTCGCCTCCAGAGTGTACTTCCAACAGAGATAGACTACTTGAAGACAAG 1466  
| | | | |  
Db 1517 ATAGAGCTTANGCCCCCGGAATGTTACTTCTCAGAAAATTTACTTGAAGACAAG 1576  
QY 1467 GTGCAAGCAACTTTGTAGGAGAGAGACAATGAAGAGAGAGATGAAGATTCAAG 1526  
| | | | |  
Db 1577 GTTCAAACTCATTTGTGAAGAACGCCGCGGANNNNNNNNNNNNNNNNNNNNNNNNCAAA 1636  
QY 1527 GTGAGATCAATGCTTGTAGTCCAAAGCCCCAGAAAGTCTCTGAAGAAAGATGACAAATG 1586  
| | | | |  
Db 1637 GTTCGTATCAATGCTCTGTAGCCAAAGCAANNNNNGTTCCCGAGAGGGATGATCATG 1696  
QY 1587 CAAGATGAACCCCTGGCTGGAACAATGTTCTGTATCATCTGGAATGATTCAAGTTC 1646  
| | | | |  
Db 1697 CAAGATGTACACCTTGGCTGGAAACAATACTAGGACCATCTGGAATGATTCAAGTTC 1756  
QY 1647 TTCTTGGCCAAAGCGAGGCTTGAATGTAGGGAATGAACCTGCCACGATTGTTTAT 1706

Db 1757 TTCCTGGTCAAGTGGAGGGCTTGACGTTGAAGGCAATGAACCTTCCCTGTTGTTTAT 1816  
QY 1707 GTTCTAGAGAGAAACGACGAGGCTATTAACCATCATTAAGAAAGCTGTGCTATGAATGCA 1766  
| | | | |  
Db 1817 GTGTCTCGTGNNNNNCGTCTTGATTTCCACATCAACAAGAGGCTGTGCCATGAATGCA 1876  
QY 1767 TTGTCGCAAGTCTCTGCTGTACTAACAATGCTCCATATTGTTAACTTGATGTGAT 1826  
| | | | |  
Db 1877 CTGTTCGTATACAGCTGTCTTACTTAATGGGCAATACATGTGAATCTTGATTTGAC 1936  
QY 1827 CACTACATCAACAACAGCAAGGCTATTAAGAAAGCAATGTGTTTATGATGACCCCTTTA 1886  
| | | | |  
Db 1937 CACTACATCAATTAAGCAAGGCTCTTGAGAAAGCTATGTCTTCTTATGAGACCAAAAC 1996  
QY 1887 CTAGAAAGAGGTTGCTATGTACAGTTCCTCAAGATTGATGGGATGATCGCCAT 1946  
| | | | |  
Db 1997 CTAGGAAGAAATGTCTGTATGTCCAAATTCCTCAGAGTTGATGTTGATAGGAAT 2056  
QY 1947 GACCGATATGTAACCGGAATGTGTCTTTTGATATCAACATGAAGTTTGATGT 2006  
| | | | |  
Db 2057 GACCGATATGCAACAGGAACACACTGTGNNNNNCGATATTAACCTTGAGAGTCTTGACGGC 2116  
QY 2007 ATTCAGGCTCCAAATTTATGTTGTACTGTGATGTATTTAGAAGGCAAGGATTAATGT 2066  
| | | | |  
Db 2117 ATTCAGGCGCCAGTTTATGTGGAACTGTGTGTGTTTAAACAAGACGCGCTTATATGT 2176  
QY 2067 TATGATGCCCCCAAAACAAGAACCAATCAAGAAAGCTTGCAACTGTGCCCCAAGTGG 2126  
| | | | |  
Db 2177 TATGAGCCTCCAGTCAAGNNNNNNNGCCA----- 2206  
QY 2127 TGCTTTTGCTGTGCTGCTTTGGCAATAGGAACAAGAAAGACTACCAAAACCAAAACA 2186  
| | | | |  
Db 2207 GGCTTCTTCTCTCGCTTGTNNNNNAAGNNNNNGACGTCAAATCTAAGAAAGCTCG 2266  
QY 2187 GAGAAGAAAAGTTATTTTTCAGAGAAGAGAACAACCAATCCCTGCAATATGCTCTT 2246  
| | | | |  
Db 2267 GAAAAGAAAGTCA-----TAGACAGCAGACAGTCTGTACCAAGATTTAATCTC 2320  
QY 2247 GSTGAATTGACGAAGCTGCTCCAGAGC-----TGAGATGAAGAGCGCGTATTTGTA 2300  
| | | | |  
Db 2321 GAAGATATAGGAAGGAGATTGAAGTCTCAGTTGATGATGAGAAATCGCTGATTAATG 2380  
QY 2301 AATCAACAAAATTAGAAAAGAAATTTGGCCAACTCTCTGTTTGTATACCAACTT 2360  
| | | | |  
Db 2381 TCTCAATGAGCTTGGAAGAGAGATTGGCCAGTCCAGTGNMNNNGTAGCTCTACTCTG 2440  
QY 2361 CTCGAGATGTGAACCTTGAAGAGTCAAGTCTGCTTCTTTTGAAGAAAGCTATA 2420  
| | | | |  
Db 2441 ATGAATATGTTGTTCCACAATCTGCAACTCCNNNNNNNNNNNNNNNNNNNNNTTATT 2500  
QY 2421 CATGTCATTAGTTGTGTATGAAGACAAGACAGACTGGGAAAGAGATTGCTGATC 2480  
| | | | |  
Db 2501 CATGTCATCAGCTGTGCTATAGAGACAACAACCTGCGGAACCTGAGATTGGGTGATC 2560  
QY 2481 TATGATCAGTTACAGAGATATTCTAATGTTTCAAGATGATTTGTCATGTTGCGG 2540  
| | | | |  
Db 2561 TATGTTCTGTACAGAGACATTCTCACCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2620  
QY 2541 TCAATTTACTGCATACCTTAACGCGGTGCAATTAAGGTTCTGCACTCTGAATCTTTCA 2600  
| | | | |  
Db 2621 NNN 2680  
QY 2601 GATGCTCTTCAACAGGTGCTGCTGGGCTCTTGGGCTATTTGAGATCTTTCAGCAAT 2660  
| | | | |  
Db 2681 GATGCTTGAATCAAGTGTTCGGTGGGCTCTTGGTTCCATTTGAATCTTTTCAGCAGG 2740  
QY 2661 CATGCTCTTGTGTAAGGTTATGTTGTTGCGGTGGAATTTTGAAGATTTTCTCTAC 2720  
| | | | |  
Db 2741 CATGTCCCATATGTTAGTGTATGAGAGCGCGCTTAAATTCCTGAGAGATTTGCTTAT 2800  
QY 2721 ATCACTCCATCGTGTATCTTGACATCTATTCCTCTTGCTTACTGTACATTTGCCCT 2780  
| | | | |

Db 2801 ATCAACACAATTTATCCACTCATCAATCCCGCTCCTCTGTACTGATATTGCCA 2860  
QY 2781 GCCATCTGTTATTTGACAGGAAATTTATCACTCCAGAGCTGAATATGTTGCCAGCTG 2840  
Db 2861 GCAGTTGTCTTCTCACTGGGAAGTTCATCAATCCCAAGATTAGTAACCTAGAGAGTGT 2920  
QY 2841 TGGTTCATGCACTTTTATCTGCAATTTTGTCTACGAGCATCTTGAATAGAGATGAGT 2900  
Db 2921 TGGTTATATCGCTCTTATCTCAATCTTTCACACTGTGTATCTTGNNNNNNNNNNN 2980  
QY 2901 GGTGTGAATGATGACTGTGAGAGAAAGCAAGTCTGTGATGATGAGTGTGTC 2960  
Db 2981 NNN 3040  
QY 2961 TCACACCTCTTGTCTGTGTTCCAGGACTTCTCAAGTCATAGCTGTGTGATACAGC 3020  
Db 3041 GCGCATTTATTTGCCGCTCTCCAGGCTCTCTGAAGGTGCTGTGTATCGACACGAGC 3100  
QY 3021 TTCACCGTGAATCAAGG--GTGAGATGATGAGAGTCTCAAGCTATATACATTC 3077  
Db 3101 TTCACTGTCACTCTAAGGCCACTGACGAAAGGTGATTTGCGAGCTCTACATGTTTC 3160  
QY 3078 AAATGACTAATCTTATGATACCTCTTACCACTTGTCTTATGAACTTCAATGTTGTTG 3137  
Db 3161 AAGTGACAAGCTTCTGTATCCCAACCACTATTTGATCATCACTGTGCGCGTG 3220  
QY 3138 GTGCTGGCTTCAATGCGATCAATACGATATGATGATGAGGCGCCCTCTTGGG 3197  
Db 3221 GTGCTGGCAATTCCTACGCAATCAATAGGGTTACAGTCAATGGGACCTTTTCGGG 3280  
QY 3198 AAGCTATCTTTGCAATTTGGGTGATGTTGTCATCTTATCCCTTTCAGAGTTGTT 3257  
Db 3281 AAGCTCTCTTTGCGTTCTGGGTGATGTTGTCACCTGTACCCCTTCTCAAGGCGCTCATG 3340  
QY 3258 GGAAGCAAAACAGACACCAAGATGTCATGCTGTGTCATCTGTGCTGCTCAATC 3317  
Db 3341 GGAAGCAAGAACCCGACCGCCGATGTTGTTGTTGCTGTGCTATCTCTTGGCTGATC 3400  
QY 3318 TTCTCGCTCTTTGGGTTCGATGATCTTCTCTTGGCAAGATGATGCTCGCTTCT 3377  
Db 3401 TTTTCCCTGATGTGGGTTCGATGATCAATCAACACCGGGTCACTGTCTGATATC 3460  
QY 3378 GAGAGTGTGTTGATTGCAACTAGATGCA 3411  
Db 3461 GCGAATGTGCATCAACTGCTAGATGAGCTGA 3494

RESULT 7  
AY112236  
LOCUS 3728 bp mRNA linear HTC 17-OCT-2002  
DEFINITION Zea mays CU1160\_1 mRNA sequence.  
ACCESSION AY112236  
VERSION AY112236.1 GI:21216826  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 3728)  
AUTHORS Coe, E.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
COMMENT  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSL, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,

www.tigr.org; or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).  
FEATURES  
Source  
1.3728  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:630049"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"  
BASE COUNT 782 a 845 c 925 g 878 t 298 others  
ORIGIN  
Query Match 30.7%; Score 1209.6; DB 11; Length 3728;  
Best Local Similarity 59.9%; Pred. No. 8.7e-260;  
Matches 1963; Conservative 0; Mismatches 1210; Indels 102; Gaps 7;  
QY 210 GACGCGATCCCGGCGGAGCGCGCGGAGCAGAACGGGCAAGTGTGCCAGATTGC 269  
Db 244 GCGACGCGGACGCGCGTGAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTCCAGATGTC 303  
QY 270 GCGCAGACGTCGCGCTTGGCCCGCGGCGGAGCCCTTGTGCGTGCAGACGATGCGCC 329  
Db 304 GCGCATGCGGTGGCACTACGCGCGGAGGAGACGTCTTCAACCGCTGCGAGCTGCGGG 363  
QY 330 TTCCCGTCTCGCGGACTGTCTAGATACGAGCGCGGAGGAGGACGAGAACTGCCCC 389  
Db 364 TTCCCGGTGTGCGCGCCCTGTCTAGATACGAGCGGAGGAGGAGGAGGAGGAGGAGG 423  
QY 390 CAGTGAAGACTCGATACAGCGCTCAAGGCGTGCACGCTGTGACCGGTGACGAGAG 449  
Db 424 NAGTGNNN 483  
QY 450 GAGACGCGCTGATGACCTGACCAACGAGTTCACT-----GGACGGCAT 497  
Db 484 GACGATACTGATGCCGATGATGCTAGCACTTCACTTCACTGCTGCAATGACGAC 543  
QY 498 GACTGCACTCTGTGCGCGAGTCTCTTACGCGGACATGAGCTACGCGCGTGAAGT 557  
Db 544 CAGAAGCAAGATGTCTGACAGATGCGCAGCTGCGCATGAATGCTNNNNNNCAGCGG 603  
QY 558 GACCTAATGCGCGGCAAGCTTTCAGCTCAACCCCAATGTTCCACTCCTCACCAAC 617  
Db 604 GATGTTGCGCGCCCAAGTATGACAGTGTGAGATCGGGTTACCAAGTACGACAGTGT 663  
QY 618 GGGCAATGTGATGATGATCCACCGGAGCAGCAGCGCTGTGCTTTCATGGGT 677  
Db 664 GAGATCCCTCGGGATACATCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 723  
QY 678 GGTGGGGAAGAGATACATCCCT---TCCTTATGCGGATCCCACTTACCTGTGCA 734  
Db 724 GTTGTCTCCCTGACCATCATATGATGTCTCTTACTGGAACATTGGCAGGCGCGCCA 783  
QY 735 CCCAGTCTATGACCCCATCAAGATCTTGTGCTGATATGGGTATGTAG-----T 785  
Db 784 TTTCCTATATGATCATTCATCAATCCGTGAGGGAATTTCTGTAGCGTTGGGAAT 843  
QY 786 GTTCTTGAAGGAACGATGAGAATTGGAAGCAGACAAG----- 829  
Db 844 GTTGCCTGGAAGAGAGGTGATGCTGGAATGAAGCAGACAAGGAATTTCCC 903  
QY 830 -----GAGATGCAACGAGCGGGAATGATGTGTGTGATGATGTG-- 873  
Db 904 ATGACGAATGCAACGATTTGCTCCTCTGAGGGCGCGGGTGTGTGATATTTATGCA 963



QY 874 -----ACGATGCTGATCTACCACTAAATGATGAAGCAAGACAACACTGTCC 920  
Db 964 TCAACTGATTACAACATGGAGAAGATGCCCTTATTAACGATGAAGAACTCGCCAGCCTCTATCT 1023  
QY 921 AGGAAAATTCACCTTCCATCAAGCCAGATTATCCATATAGATGATTATCATTTATTCGG 980  
Db 1024 AGGAAAGTTCACCTTCTCTCCAGGATTAATCCATACAGGATGTGCTATGTGCTACGA 1083  
QY 981 CTGTGTTTGGGGTCTTCTTCCACTACCGAGTGTGATCCGGTGAATGATGATTT 1040  
Db 1084 TTGATGTTCTAAGCATCTTCTTGCACTACCGGATCACAATCTGTGCGTAATGATAC 1143  
QY 1041 GCTTGTGGCTCATATCTGTTATCTGTAATCTGTTTGCCATGTCTTGATTTGAT 1100  
Db 1144 CCACTGTGCTTCTATCTGTATATGTGAGATCTGTTGTGCTCTTCCGTGATATGAT 1203  
QY 1101 CAATTCACCAAGTGTCCCTATTTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGG 1160  
Db 1204 CAGTTTCCAAAGTGTTCATCAATCAACCGGAGACTTACTTGATAGACTCGCATTAAGG 1263  
QY 1161 TTGACAAGAGAGCGCACCATCTCAACTGTCTCAATGATTTCTTGTCAGTACGGTT 1220  
Db 1264 TATGACCGGGAAGGTAGCCATCTCAGTTGCTGTGACANNNNNGTCACTACTGTC 1323  
QY 1221 GATCCCTTAAAGAACCTCTTGTGCTCACACAATACTGTCTATCTATCCTTTCGGTG 1280  
Db 1324 GACCCAATGAAGGACCTCTCTTGTCTCACTGCCAATACCGTGTATCCATTCTCGCTGTG 1383  
QY 1281 GATTATCTGTGATTAAGTTTCTGCTATGTTTCTGATGATGCTGCAATGCTTAACG 1340  
Db 1384 GACTATCCTGTGATTAAGTCTCTGTCTATGATGATGAGCTGTATGCTGACA 1443  
QY 1341 TTTGAAGCATTTATCTGAACAATCTGAATTTGCAAAAGAAATGGGTTCTTCTGCAACGG 1400  
Db 1444 TTGATGCACTAGCTGAGACTTCAAGATTGCTAGAAAATGGGTCATTGTTAAGAG 1503  
QY 1401 TACAATATTGAACCTCGCGCTCCAGAGTGTACTTCCAAGCAAGATAGACTACTTGA 1460  
Db 1504 TACAACATTTGAACCTAGAGCTCTCGAATGCTACTTCTCCAGAAAATTTGATTTACTGAAG 1563  
QY 1461 GACAAGGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAAGATATGAGGA 1520  
Db 1564 GACAAAGTGCAACCTTCATTGTTTAAAGACCGCGGNNNNNNNNNNNNNNNNNNNN 1623  
QY 1521 TTCAAGGTGGAATCAATGCTTAGTTGCCAAAAGCCAGAAAATTCTGAAGAAGATGG 1580  
Db 1624 NNCAAAATTAAGGTAAATGGCTTGTGCTAAGGACANNNNNGTCCCTGAGGAAGATGG 1683  
QY 1581 ACAATGCAAGATGGAACCCCTGGCTGGAACAATGTTGCTGATCATCTCGAATGAT 1640  
Db 1684 ATCATGCAAGATGGCACAACCATGGCCAGGAAAACAATACAGGACCATCTCGAATGAT 1743  
QY 1641 CAGGTCTCTTGGCCAAAGCGGAGCCTTGAAGTGTGAGGGAATGAAGTGCACGATTTG 1700  
Db 1744 CAGGTTTCTTGGTCAAGTGTGCTTGTGATACTGAGGTAATGAGCTAACCCGTTTG 1803  
QY 1701 GTTATGTTTCTAGAGAGAAAGAGAGCTATTAACCATCATAGAAGAGCTGTGCTATG 1760  
Db 1804 GTCTATGTTTCTCGTGMNNNNCGTCTCGAATTTCCAGCATCAAGAAAAGCTGTGCCATG 1863  
QY 1761 AATGATTTGGTCCGAGTCTCTGCTGTACTTAACAATGCTCCATATTTGTTAACTTGAT 1820  
Db 1864 AATGCTTTGTCGCGCTCAGCTGTGCTTAACAATGACAATACATGTTGAATCTTGAT 1923  
QY 1821 TGTGATCACTACATCAACAACAGCAAGGCTATTAAGGAAGCAATGTGTTTATGATGAC 1880  
Db 1924 TGTGATCACTACATCAACAACAGTAAGGCTCAGGGAAGCTATGTGCTTCTTATGAT 1983  
QY 1881 CCTTACTAGGAAGAAGGTTTGTATGTACAGTCCCTCAAAAGATTGATGGATTGAT 1940  
Db 1984 CCTAACCTAGGAAGAGTGTCTGTATGTTCAAGTTCCCGAGAGGTTGATGATGAT 2043  
QY 1941 CGCCATGACGATATGCTAACCGGAATGTGTCTTTTGTATATCAACATGAAGGTTTG 2000

Db 2044 AGGAATGATCGATATGCCACAGAGAACCCGTGNNNNNGATATTAATTGAGAGTCTT 2103  
QY 2001 GATGATATTCAGGGTCCCAATTTATGTGTGTACTGATGTGTATTTAGAAGCGAGCATTA 2060  
Db 2104 GATGCAATCAAGAACAGGTTTATGTGGGCACTGCTGTGTTTCAACAGAACAGCTCTA 2163  
QY 2061 TATGTTATGATGCCCCCAAAACAAGAGCCACCATCAAGACTTGCAACTGTGGCCC 2120  
Db 2164 TATGTTATGAGNNNNNAATTAAGCAAAG----- 2193  
QY 2121 AAGTGTGCTTTTCTGCTGTGCTGCTTGGCAATAGGAAGCAAAAGAACTAACCAACCC 2180  
Db 2194 AAGGTGTTCTTGTCTATCACTATGTGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2253  
QY 2181 AAACAGAGAGAAAAGTTATTTTCAAGAAAAGAGAACCAATCCCTGCATAT 2240  
Db 2254 GGCTCAGACAAAGNN 2313  
QY 2241 GCTCTGTGTAATTTGACGAAGCTGCTCCAGAGCTGAGATGAAGCCGGTATTTGA 2300  
Db 2314 NNNNNNTAGAGAGAGGAGTTGAAGCCGCTGATTTGATGATGAGAAATCACTTCTTATG 2373  
QY 2301 AATCAACAAAATTTAGAAAAGAAATTTGGCCAACTTCTGTTTGTATCATCCACACTT 2360  
Db 2374 TCTCAATAGCTTGGAGAGAGATTTGGCCAATCTGCAAGCTTGTGTGCTCACTCTG 2433  
QY 2361 CTCGAGATGTGGAACCTTGAAGAGTGCAAGTCTGCTTCTTTTGAAGAAAGCTATA 2420  
Db 2434 ATGGAATNNNTATC 2493  
QY 2421 CATGTCATTAAGTGTGTTATGAAGACAAGACAGACTGGGAAAAGAGATTGCTGATC 2480  
Db 2494 CATGTCATTAAGTGTGCTACGAGACAGATTTGAATGGGAAGCTGAGATTGGGTGATC 2553  
QY 2481 TATGATCACTTACAGAGATATTTCTAAGTGTTCAGATGATGATGCTGCTGGCG 2540  
Db 2554 TATGTTCTGTGACGGAAGATATTTCTCACTGGGTTCAAGATGACGACGAGGCTGGCG 2613  
QY 2541 TCAATTTACTGCATACCTTAAACGGGTGCTGATTCAAAGGTTCTGCACCTCTGAATCTTCA 2600  
Db 2614 TCGATCTACTGATGCTTAAAGCGCGGCTTCAAGGATCGGCTCCCATCAATCTCTCA 2673  
QY 2601 GATGCTCTTCAACAGGCTTCCGTGGGCTCTTGGGCTATTTGAGATCTTTCAGCAAT 2660  
Db 2674 GACGCTGAAACAGGCTCTCCGCTGGGCTCTCGGTTCAAGTGAATCTTTCAGCCGG 2733  
QY 2661 CATGCCCTCTTGTGTTAGGTATGTGCGGCTGAATTTTGAAGAATTTTCTTAC 2720  
Db 2734 CATGNNNNNTATGTACGGGTACGAGAGACGCTGAAGTTCTTGAGAGATTCGCTTAC 2793  
QY 2721 ATCAACTCCATCGTGTATCTTGGACATTAATTCCTCTTGGCTTACTGTACATTCCT 2780  
Db 2794 ATCAACACCAACCATCTACCCGCTCACGTCCTCCGCTCTCATTTACTGTATCTGCT 2853  
QY 2781 GCCATCTGTTATTTGACAGGAAATTTATCACTCCAGAGCTGAATATGTTGCCAGCTG 2840  
Db 2854 GCCATCTGCTGCTCACGGGGAAGTTCAATCATCCAGAGATCAACAACTTCGCTATATC 2913  
QY 2841 TGGTTCATGTCATTTTATCTGCAATTTTGTCTACAGACATCTAGAAATGAGATGAGT 2900  
Db 2914 TGGTTCATCTCTCTTCAATCTGATCTTCCACGGGTATCTGAGATGAGGTGAGC 2973  
QY 2901 GGTTTGAATTTGATGATGCTGTGAGGAATGAGCAAGTTCTGGTCAATGAGGTGTCTC 2960  
Db 2974 GGCGTGGCATGACGAGTGTGAGGAACGAGCAAGTTCTGGGTCACTGAGGCAATCTCC 3033  
QY 2961 TCACACTCTTTGCTGTGTTCCAGGAGCTTCTCAAGGTATAGCTGTGTTGATPACAAGC 3020  
Db 3034 GCCACCTCTTCCGCTCTTCCAGGGCTCTCAAGTGTGCTCCGCGCATGACACCAAC 3093  
QY 3021 TTCACCGTGACATCAAGGCTGAGATGATGA---GGAGTTCTCAGAGCTATATACATTC 3077

Db 3094 TTCAACGTCACCTCCAGGCGCTCGATGAAGACGGGACCTTCGGGAGCTGTACATGTTTC 3153  
QY 3078 AAATGGACTACCTTATGTATACCTCCCTAACCACTTGCTTCTATTTGAATTCATTTGTTGTC 3137  
Db 3154 AAGTGGACGACACTTCTGATCCCGCCCAACCACTCCTGATCATCAACCTGTCGGCGTT 3213  
QY 3138 GTCGCTGGCGCTTCAAAATGCGATCAATAACGATATGATCATGGGGCCCCCTTTTGGG 3197  
Db 3214 GTTGCCGCATCTCTACGCGCATCAACGGGTACAGTCGTGGGGTCCGCTTTCCGC 3273  
QY 3198 AAGCTATCTTTGCAATTTGGGTGATGTCATCTTATCCCTTTCTCAAGGTTGTT 3257  
Db 3274 AAGCTCTTCTTGCCCTTCTGGGTGATGTTCACTGTACCCGTTCTCAAGGTTCTCATG 3333  
QY 3258 GGAAGCAAAACAGACACCAACGATTTGTCATCGTCTGTCCATTCTGCTGCTTCAATC 3317  
Db 3334 GGTCGGCAGAACCGCACCCCGACCATCGTGTGTCTGGGCGATCCTGCTGGCGTCAATC 3393  
QY 3318 TTCTCGCTCCTTTGGGTTCGATGATTCCTTTCTTCTGCGAAGATGATGTCGGCTTCTT 3377  
Db 3394 TTCTCTTCTGCTGTGGGTTCGATGATTCCTTTCAACCAACCGCGTCACTGCCCCGATACT 3453  
QY 3378 GAGAGTGTGTTGGATTGCACTAGGATGTCAG 3412  
Db 3454 CGAAGCTGTGCATCAACTGCTAGGAGGTGGAAG 3488

RESULT 8  
AY104236  
LOCUS AY104236 2872 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PC0121439 mRNA sequence.  
ACCESSION AY104236  
VERSION AY104236.1 GI:21207314  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
REFERENCE 1 (bases 1 to 2872)  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 2872)  
AUTHORS Coe,E.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

FEATURES  
source location/Qualifiers  
1..2872  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:637064"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"  
BASE COUNT 730 a 600 c 688 g 852 t 2 others  
ORIGIN

Query Match 29.5%; Score 1161.6; DB 11; Length 2872;  
Best Local Similarity 68.4%; Pred. No. 4.7e-249;  
Matches 1679; Conservative 1; Mismatches 740; Indels 34; Gaps 4;  
QY 903 GCAAGACAACTGTCCAGGAATAATTCACCTTCATCAAGCCAGATTAATCCATATAGG 962  
Db 13 GCACGGCTACCTTAAGTCGATAGTCCGATATCTCCAAACGAGCTTAACCTTATCCG 72  
QY 963 ATGATTTATCATTTATCGGCTTGTGTTTGGGGTTCTTCTCCACTACCGAGTATGCAT 1022  
Db 73 ATCGTATTTGTTCTCCGGCTTATCATCTATGTTTCTTCTTCAATATCGTATTACTCAT 132  
QY 1023 CCGGTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGTGAATCTGTTTGGC 1082  
Db 133 CCAGTGAAGATGCTTATGGGTGTGCTGTATCTGTATTTGTGAAGTTGGTTTGGC 192  
QY 1083 ATGCTTGATTTCTTGATCAATTCACCAAGTGTTCCTTATTTGAGAGAGACTTAACCTA 1142  
Db 193 TTGTTCTTGCTTCTAGATCAATTCACCAAGTGTATCTTATCAACCGTGAACCTTACCTC 252  
QY 1143 GACCGGCTGTCACTGAGGTTGCAAGGAAGGCCAGCATCTCACT-TGCTCAATTGA 1201  
Db 253 GATAGACTTGCAATTGAGATATGATAGGAGGGTGAGCCATCCCAAGTTGGGCTCCAATCGA 312  
QY 1202 TTTCTTTGTCACTACGCTTGATCCCTTAAGGAACCTTCTTGTGTCACAACAATACTGT 1261  
Db 313 TGCTTTTGTAGTACAGTGTGATCCACTTAAGGAACCTTCTTAATTACTGGCAACACTGT 372  
QY 1262 TCTATCTATCTTTCGTTGGATTTATCTGTTGATTAAGGTTTCTGTATGTTTCTGATGA 1321  
Db 373 CTTGTCATTTCTGCTGTGATTAACCTGTTGACAAAGTATCATGTTATGTTTCTGATGA 432  
QY 1322 TGCTGTCGAATGCTTAACGTTTGAAGCATTTATCTGAACATCTGAATTTGCAGGAATG 1381  
Db 433 CGGTTCACTATGTTGACTTTTGAAGCGCTATCTGAACCGAGAGTTTGCAAGGAATG 492  
QY 1382 GGTTCCTTTCTGCAACGCTACAATATTGAACCTTCGCTCCAGAGTGTACTTCCACA 1441  
Db 493 GGTTCCTTTTGAAGAAACACATATTTGAACCTGAGGCTCCAGAGTTTACTTTGCTCG 552  
QY 1442 GAAGATAGTACTTGAAGACAAGGTGCGAGCAAACTTTGTTAGGAGAGAGAGCAAT 1501  
Db 553 AAAGATAGATTACCTTAAGACAATAATACACCTTCTTTGTGAAGAAAGCGGCTAT 612  
QY 1502 GAAGAGAGATATGAGAAATTCAGGTGAGATCAATGCTTAGTTGCCAAAGCCAGAA 1561  
Db 613 GAAGAGGAGTGAAGAGTTCAAGATGAGATGCTTGTGCAAAAGCCGCAAAA 672  
QY 1562 AGTTCCTGAAGAGATGACAATGCAAGATGAACCCCTGGCTGGAACAAATGTTG 1621  
Db 673 AATACTGAGAGGGCTGACCATGCTGATGGCACTCCTTGCTGGGAATTAACCTAG 732  
QY 1622 TGATCATCTGGAATGATTCAGGCTTCTTGGCCAAAGCGAGGCTTGACTGTGAGGG 1681  
Db 733 AGATCATCCMGATATTCAGATATTCTTGGGCCACAGTGTGGCTTGACACGGATGG 792  
QY 1682 AAATGAACCTGCCACGATGTTTATGTTTCTAGAGAAAAGCAAGCAAGCTATTAACATCA 1741  
Db 793 GAATGAGTTGCCACGGCTTGTATGTTTCTGTTAAAGAGCCAGGCTTCCAGACCA 852  
QY 1742 TAAGAAAGCTGGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1801  
Db 853 CAAGAAAGCTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912  
QY 1802 ATATTGTTAACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1861  
Db 913 TTATCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972  
QY 1862 AATGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1921  
Db 973 TATGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1032

```
QY 1922 AAGATTGATGGATGATCGCCATGACCGATATGCTAACCGAATGTTGCTTTTGA 1981
    |||||
Db 1033 AAGATTGATGGATATAGACTTCATGATGATATGCAACCGAATGTTGCTTTTGA 1092
QY 1982 TATCAACATGAAGGTTGGATGCTATTCAGGGTCCAAATTTATGTTGACTGATGCT 2041
    |||||
Db 1093 TATTAATATGAAGGCTTAGATGGCAATTCAGGACCTGTTTATGTTGGAAACAGATGCTG 1152
QY 2042 ATTTAGAGGCAAGCATATATGTTATGATGCCCCCAAAACAAGAGCCACCATCAAG 2101
    |||||
Db 1153 TTCAATAGGCAAGCCTTGATGCTATGAT-----CCTGATTTGAC 1194
QY 2102 GACTTGCAACTGCTGACCAAGTGCTTTTGTGCTGCTTTGGCAATAGAGCA 2161
    |||||
Db 1195 AGAAGCTGATTTGGAGCCTTAATATCATTTAAAGTTGCTGCGGAGAGAAAAGAA 1254
QY 2162 AAAGAAGACTACCAACCCAAAACAGAGAAGAAAAGTTATTTTCAAGAAAGAA 2221
    |||||
Db 1255 GGACAAGAGCTATATGATTTCCAAAAACCGTATA-----TGAAGAGAACAGA 1302
QY 2222 GAACCAATCCCCGTCATATGCTCTTGTTGTAATGACGAGCTGCTCCAGAGCTGAGA 2281
    |||||
Db 1303 ATCTTGGGCTCCCATCTTCAACATGGAAGATATAGAAGAGGATTTGAAGGTTACGAGA 1362
QY 2282 TGAAGAGGCGGCTATTGTAATCAACAAAATTAGAAAAGAAATTTGCCAATCTTCTGT 2341
    |||||
Db 1363 TGAAGGCTCACTGCTTATGCTCAGAGAAGCTTGAGAAACGCTTTGGCAGTCTCAAT 1422
QY 2342 TTTTGTATACCCACACTTCTCGAAGATGTTGGAACCTTGAAGAGTGCAGTCTGCTTC 2401
    |||||
Db 1423 TTTTATTGATCCACCTTATGATCAAGGTGCATACCCCTTCAACAAACCGAGTTC 1482
QY 2402 TCTTTGAAAAGAGCTATACATGTCATTAAGTTGTTATGAGAACAAGACAGACTGGG 2461
    |||||
Db 1483 CCTGCTAAAGAGAGCTATACATGTCATTAAGTTGTTATGAGATAAACAGAAATGGG 1542
QY 2462 AAAGAGATTTGGCTGATCTATGATCACTTACAGAGATATTTCTACTGTTCAAGAT 2521
    |||||
Db 1543 GAAAGAGATCGGATGATATATGCTCTGTTACTGAAGATATTTTAACTGTTCAAGAT 1602
QY 2522 GCATTGTCATGCTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATCAAGGTT 2581
    |||||
Db 1603 GCATGCAAGAGGTTGATATCACTGATGCACTTGGCCTTCTTCAAGGTTTC 1662
QY 2582 TGCACCTCTGAATCTTTACAGATGCTCTCACCAAGTGCTTGGGCTCTGCTAT 2641
    |||||
Db 1663 TGCTCCGATTAATCTTCTGATGCTCAACCAAGTTAGCTGGGCTCTTGTTCAGT 1722
QY 2642 TGAGATCTTCTCAGCAATCATGCCCCCTTTGGTATGGGTATGGCGGCTGAAT 2701
    |||||
Db 1723 TGAATTTCTACTTAGCAGACATGTCCTATCTGATGTTACATGAAGGCTAAAGCT 1782
QY 2702 TTTGGAAGATTTTCTCTACATCACTCCATGCTGATCTTGGACATCTATCCCTCT 2761
    |||||
Db 1783 TCTGGAGAGACTGGCATACATCAACCAATGTTTATCCAAATTAATCATCTATCCACTAGT 1842
QY 2762 GGCTTACTGTACATTTGCTGCCATCTGTTTATTTGACAGGGAATTTATCACTCCAGACT 2821
    |||||
Db 1843 AGCATACTGCGTCTTCTGCTATCTGTTTACTCACCAACAAATTTATTAATCTCCTGCT 1902
QY 2822 GAATTAATGTTGCCAGCCTGTTGATGTCATCTTTTATCTGCAATTTTGTCTACGAGCAT 2881
    |||||
Db 1903 TAGCAATTAATGCTGGGCGCTTCTTCACTGCTTTTGTCTTCCATCTTCCACTGCTAT 1962
QY 2882 CCTAGAAATGAGATGAGTGTGTTGGAATGATGACTGTGAGGAATGAGCAGTTCTG 2941
    |||||
Db 1963 TTTGAGCTTCGATGAGTGTGTTGCAATGAGATGTTGTTGTTGTTGTTGTTGTTGTTG 2022
QY 2942 GGTCAATGAGGTGTCTCACACCTTTTGTGTTGTTCCAGGACTTCTCAAGTCAAT 3001
    |||||
Db 2023 GGTCAATGAGGTGTCTCACACCTTTTGTGTTGTTCCAGGACTTCTTAAAGTGTCT 2082
QY 3002 AGCTGCTGTGATACAAGCTTCAACCGTCACTCAAGG--GTGAGATGATGAGAGTT 3058
```

```
Db 2083 AGCAGGATCGACACAACTTCACCGTGCATCAAGGCAACCGATGATGATGTTT 2142
QY 3059 CTCAGAGCTATATACATTCAAAATGACTACTTATGATACCTCTACCACTGCTCT 3118
    |||||
Db 2143 TGCTAGCTGTATGTGTTCAGATGAGCAACTTCTGATCCCCCCACCACTGTGCTGT 2202
QY 3119 ATGAACCTCATGCTGTGTGCTGCTGCGCTTCAATGCGATCAATACGATATGAGTC 3178
    |||||
Db 2203 GATTAACCTGTTGGTATAGTGTGAGTGTGCTATGCTATCAACAGTGGCTACCAATC 2262
QY 3179 ATGGGCCCCCTCTTTGGGAAGCTATCTTTGCAATTTTGGTGATGTCATCTTATCC 3238
    |||||
Db 2263 ATGGGCTCCACTATTCGGGAAGCTGTTCTTTGCAATCTGGGTGATCTCCACTTACC 2322
QY 3239 CTTCTCAAGGTTGTTGTTGGAAGGCAAAACAGACACCAACGATTTGATCTGCTGTC 3298
    |||||
Db 2323 TTTCTGAAGGTTCTCATGCGGAAGCAGAACCGCACACCGACCATGCTCATTTGTTGTC 2382
QY 3299 CATTCGCTGCTTCAATCTTCTGCTCTTTGGGTTGGATTGATCTTCTCT 3352
    |||||
Db 2383 CGTCTTCTTGTTCATTAATTCGCTGCTGTGTGGTGAAGATGACCCCTTCAT 2436
```

```
RESULT 9
CC379931/c 877 bp DNA linear GSS 19-MAY-2003
LOCUS
DEFINITION
PUHOP50TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM5BTa498J04,
genomic survey sequence.
ACCESSION
CC379931 GI:30855570
VERSION
CC379931.1
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 877)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
location/Qualifiers
Source
1..877
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BTa498J04"
/clone_1lb="ZM_0.6_1.0_KB"
/notes="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
BASE COUNT
290 a 220 c 174 g 193 t
ORIGIN
Query Match 20.6%; Score 811; DB 29; Length 877;
Best Local Similarity 99.9%; Pred. No. 1.3e-170;
Matches 822; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2818 AGCTGAATATATGTTGCCAGCCTGTGCTCATGTCATTTTATCTGCATTTTCTACGA 2877
    |||||
Db 822 AGCTGAATATATGTTGCCAGCCTGTGCTCATGTCATTTTATCTGCA-TTTTCTACGA 764
    |||||
QY 2878 GCATCTAGAAATGAGATGAGTGTGTTGGAATGATGACTGTGAGGAATGACAGT 2937
    |||||
```



Db 763 GCATCCTAGAAATGAGATGAGTGGTGTGGAATGATGACTGGTGAGGAATGAGCAGT 704  
QY 2938 TCTGGGTCAATGGAGGTGTGTCTCCACACCTCTTTGCTGTGTCCAGGACTTCTCAGG 2997  
Db 703 TCTGGGTCAATGGAGGTGTGTCTCCACACCTCTTTGCTGTGTCCAGGACTTCTCAGG 644  
QY 2998 TCATAGCTGGTGTGATACAAAGCTTCCAGGTGACATCAAAAGGTGAGATGATGAGAGT 3057  
Db 643 TCATAGCTGGTGTGATACAAAGCTTCCAGGTGACATCAAAAGGTGAGATGATGAGAGT 584  
QY 3058 TCTCAGAGCTATATACATTCAAATGAGACTTATGTATACCTCCTACCACTTGCTTC 3117  
Db 583 TCTCAGAGCTATATACATTCAAATGAGACTTATGTATACCTCCTACCACTTGCTTC 524  
QY 3118 TATGAACTTCATTTGGTGTGTGCTGCGCTTCAATGCGATCAATTAACGATATAGT 3177  
Db 523 TATGAACTTCATTTGGTGTGTGCTGCGCTTCAATGCGATCAATTAACGATATAGT 464  
QY 3178 CATGGGGCCCCCTCTTTGGGAAGCTATCTTGCATTTTGGGTGATGTCCATCTTATC 3237  
Db 463 CATGGGGCCCCCTCTTTGGGAAGCTATCTTGCATTTTGGGTGATGTCCATCTTATC 404  
QY 3238 CCTTCTCAAAGTTTGGTGTGGAAGGCAAAACAGACACCAACGATGTCTGTCTGT 3297  
Db 403 CCTTCTCAAAGTTTGGTGTGGAAGGCAAAACAGACACCAACGATGTCTGTCTGT 344  
QY 3298 CCATTCGTGCTGCTCAATCTTCTGCTGCTTGGGTGCGATGATCTTCTTCCGCA 3357  
Db 343 CCATTCGTGCTGCTCAATCTTCTGCTGCTTGGGTGCGATGATCTTCTTCCGCA 284  
QY 3358 AGGATGATGCTCCGCTTCTTGAAGAGTGTGATGCAACGATGATGATGATGATGAT 3417  
Db 283 AGGATGATGCTCCGCTTCTTGAAGAGTGTGATGCAACGATGATGATGATGATGAT 224  
QY 3418 CAGCTCCCCCAATCTGCATATGCTGAAGTATTTCTGCTGTGTCTTCCCATATTCAG 3477  
Db 223 CAGCTCCCCCAATCTGCATATGCTGAAGTATTTCTGCTGTGTCTTCCCATATTCAG 164  
QY 3478 TGTCTGTAGTAAAGACATGAAATGTCCCAAGTTCTTTGATCCATGCTGTAACCTACT 3537  
Db 163 TGTCTGTAGTAAAGACATGAAATGTCCCAAGTTCTTTGATCCATGCTGTAACCTACT 104  
QY 3538 TAATATCTGAGATATATCTGGGGGAAAATGAGGCTGCGCAATCTTGTGAGTGGG 3597  
Db 103 TAATATCTGAGATATATCTGGGGGAAAATGAGGCTGCGCAATCTTGTGAGTGGG 44  
QY 3598 CCGTGAATACAGCATATGCAAGTGTGATGTGACAGCATTC 3640  
Db 43 CCGTGAATACAGCATATGCAAGTGTGATGTGACAGCATTC 1  
RESULT 10  
CD441181/c 758 bp mRNA linear EST 03-JUN-2003  
LOCUS CD441181 EL01N0553C04.g Endosperm\_5 Zea mays cDNA, mRNA sequence.  
DEFINITION CD441181  
ACCESSION CD441181  
VERSION CD441181.1 GI:31356824  
KEYWORDS EST.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
REFERENCE 1 (bases 1 to 758)  
AUTHORS Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and  
Messing, J.  
TITLE Sequencing of the maize endosperm ESTs  
JOURNAL Unpublished  
COMMENT Contact: Lai, Jinsheng  
Dr. Joachim Messing's Lab  
Waksman Institute, Rutgers University  
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
Tel: 732-445-3801

Fax: 732-445-5735  
Email: jlai@waksman.rutgers.edu  
Seq primer: T7.  
FEATURES  
source location/Qualifiers  
1. 758  
/organism="Zea mays"  
/mol\_type="mRNA"  
/culivar="W22"  
/db\_xref="taxon:4577"  
/issue\_type="Endosperm of 7-23DAP"  
/clone\_id="Endosperm 5"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"  
BASE COUNT 236 a 177 c 144 g 201 t  
ORIGIN  
Query Match 18.6%; Score 731.4; DB 14; Length 758;  
Best Local Similarity 99.2%; Pred. No. 8.5e-153;  
Matches 735; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 3144 GGGCTTCAAATGCGATCAATTAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTA 3203  
Db 758 GGGCTTCAAATGCGATCAATTAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTA 699  
QY 3204 TTTCTTGCATTTTGGGTGATGTCCATCTTATCCCTTCTCAAAGTTTGGTGAAG 3263  
Db 698 TTTCTTGCATTTTGGGTGATGTCCATCTTATCCCTTCTCAAAGTTTGGTGAAG 639  
QY 3264 CAAACAGGACCAACGATGTCAATGCTGTGCTCATTCGCTGCTTCAATCTTCTCG 3323  
Db 638 CAAACAGGACCAACGATGTCAATGCTGTGCTCATTCGCTGCTTCAATCTTCTCG 579  
QY 3324 CTCTTTGGGTTGCGATGATCTTCTTCTTCCGGAAGATGATGCTCCGCTTCTTGAAG 3383  
Db 578 CTCTTTGGGTTGCGATGATCTTCTTCTTCCGGAAGATGATGCTCCGCTTCTTGAAG 519  
QY 3384 TGTGTTTGGATTGCAACTAGATGTCAGTGCATCAGCTCCCCCAATCTGCATATGCTTG 3443  
Db 518 TGTGTTTGGATTGCAACTAGATGTCAGTGCATCAGCTCCCCCAATCTGCATATGCTTG 459  
QY 3444 AAGTATATTTCTGCTGTGTGCTCCCATATTCAGTGTCTGTAGATAAGACATGAATG 3503  
Db 458 AAGTATATTTCTGCTGTGTGCTCCCATATTCAGTGTCTGTAGATAAGACATGAATG 399  
QY 3504 TCCCAAGTTCTTTGATCCATGCTGAACCTACTTAATCTGAGAGATATACTGGGGA 3563  
Db 398 TCCCAAGTTCTTTGATCCATGCTGAACCTACTTAATCTGAGAGATATACTGGGGA 339  
QY 3564 AAATGAGGCTGCGCAATCTTGTGAGTGGGCGGTGAATAACAGCATATGCAAGTGT 3623  
Db 338 AAATGAGGCTGCGCAATCTTGTGAGTGGGCGGTGAATAACAGCATATGCAAGTGT 279  
QY 3624 TTGATTTGTCAGCATCTTTATTAATTGCTGCGCAATATAGATGGGCTGAGCCGAACAGCA 3683  
Db 278 TTGATTTGTCAGCATCTTTATTAATTGCTGCGCAATATAGATGGGCTGAGCCGAACAGCA 219  
QY 3684 AGGTAATTTGATTCTGACCTGCTCCGCTGACAACTTGTTCTCAATAAGCAGCAGG 3743  
Db 218 AGGTAATTTGATTCTGACCTGCTCCGCTGACAACTTGTTCTCAATAAGCAGCAGG 159  
QY 3744 AATGCACTGCCAGTGGAACAGCAACCTGCACATTAATTAATGATGCTGTTCAATGG 3803  
Db 158 AATGCACTGCCAGTGGAACAGCAACCTGCACATTAATTAATGATGCTGTTCAATGG 99  
QY 3804 AGGGCTTGTTCATTACATGTGTCTATACTAGAAAAACAGAAATTAAGCATTAATCTA 3863  
Db 98 AGGGCTTGTTCATTACATGTGTCTATACTAGAAAAACAGAAATTAAGCATTAATCTA 39  
QY 3864 TAGTTAATTAAGTATGTAA 3884  
Db 38 TAGTTAATTAAGTAAAAAA 18

RESULT 11	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
CB633016	856 bp	mRNA	linear	EST 08-APR-2003								
OSIEB1J21.f	OSIEB	Oryza sativa (indica cultivar-group)	CDNA									
clone OSIEB1J21 5'	mRNA sequence.											
CB633016												
CB633016.1	GI:29628005											
EST.												
Oryza sativa (indica cultivar-group)												
Oryza sativa (indica cultivar-group)												
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;												
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;												
Ehrhartoideae; Oryzaceae; Oryza.												
1 (bases 1 to 856)												
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,												
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.												
Large-scale identification of ESTs involved in the interaction												
between rice and Magnaporthe grisea												
Unpublished												
Contact: Rod Wing												
Arizona Genomics Institute												
University of Arizona												
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ												
85721-0088, USA												
Tel: 520 626 3967												
Fax: 520 621 9288												
Email: http://genome.arizona.edu												
PCR Primers												
FORWARD: gta aaa cga cgg cca gtc												
BACKWARD: gga aac agc tat gac cat g												
Plate: 11 row: J column: 21												
Seq primer: gta aaa cga cgg cca gtc.												
Location/Qualifiers												
1..856												
/organism="Oryza sativa (indica cultivar-group)"												
/mol_type="mRNA"												
/cultivar="IR36"												
/db_xref="taxon:39946"												
/clone="OSIEB1J21"												
/tissue_type="Leaf"												
/dev_stage="3 week"												
/lab_host="DH10B"												
/clone_lib="OSIEB"												
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:												
XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"												
BASE COUNT	222 a	165 c	203 g	266 t								
ORIGIN												
Query Match	17.9%;	Score 704.6;	DB 14;	Length 856;								
Best Local Similarity	89.0%;	Pred. No. 8.7e-147;										

Db	302	GATGCACTGCCATGGTGTGGCGGTCAATTTACTGCATACCTAAACTACCTGCATTCAAGG	361
Qy	2579	TTCTGCACCTCTGAATCTTTCAGATCGTCTTCACCAGTGCTTCGGTGGCTCTTGGGTC	2638
Db	362	TTCTGCACCACTCAATCTTCTGTATCGTCTTCACCAAGTCTTCGGTGGCTCTTGGGTC	421
Qy	2639	TATTGAGATCTTCTTCAGCAATCATTTGCCCTCTTGGTATGGGTATGGTGGCGGTCTGAA	2698
Db	422	TGTTGAATATATCTTCAGCAACCATTTGCCCTCTTGGTATGATATGGTGGTGGACTAAA	481
Qy	2699	ATTTTGGAAAGATTTTCTCTACATCAACTCCATCGTGTATCCTTGGACATCTATCCCT	2758
Db	482	ATGTTTGGAAAGATTTTCTCTACATTAATTCATCGTGTATCCTTTACATCAATCCACT	541
Qy	2759	CTTGGCTTACTGTACATTTGCCCTGCATCTGTTTATTTGACAGGAAATTTATCACTCCAGA	2818
Db	542	CTTGGCTTACTGTACATTTGCCCTGCATCTGTTTATTTGCTAACTGGAAATTTATCACTCCAGA	601
Qy	2819	GCTGAATAATGTTGCCAGCCTGTGTTCAATCACTTTTATCTGCATTTTGTCTACGAG	2878
Db	602	GCTTACGAATGTGCCAGTCTCTGTTCAATCACTTTTATCTGCATTTTGTCTACGAG	661
Qy	2879	CATCCTAGAAATGAGATGGAGTGTGTTGGAATTGATGACTGTGTGAGGAATGACACTT	2938
Db	662	CATCCTAGAAATGAGATGGAGTGTGTTGGAATTGATGACTGTGTGAGGAATGACACTT	721
Qy	2939	CTGGGTCAATGGAGGTGTCTCTCACACCTCTTGTCTGTCTCCAGGACCTTCTCAAGGT	2998
Db	722	TTGGGTCAATGGAGGTGTCTCTCACACCTCTTGTCTGTCTCCAGGACCTTCTCAAGGT	781
Qy	2999	CATAGCTGTGTGTTGATACCAAGCTTCACCGTGACATCAAAAGGTGGAGATGATGAGAGTT	3058
Db	782	CATAGCTGTGTGATGATACCAAGCTTCACCGTGACATCAAAAGGTGGAGATGATGAGAGTT	841
Qy	3059	CTCAGAGCTATATATAC	3073
Db	842	CTCAGAGCTATATATAC	856
RESULT 12			
CB645001		859 bp	mRNA linear EST 08-APR-2003
LOCUS			
DEFINITION	OSJNEB06H07.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA		
ACCESSION	CB645001		
VERSION	CB645001.1	GI:29639992	
KEYWORDS	EST.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
AUTHORS	1 (bases 1 to 859)		
TITLE	Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.		
JOURNAL	Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea		
COMMENT	Unpublished		
	Contact: Rod Wing		
	Arizona Genomics Institute		
	University of Arizona		
	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ		
	85721-0088, USA		
	Tel: 520 626 3967		
	Fax: 520 621 9288		
	Email: http://genome.arizona.edu		
	PCR Primers		
	FORWARD: gta aaa cga cgg cca gtc		
	BACKWARD: gga aac agc tat gac cat g		
	Plate: 06 row: H column: 07		
	Seq primer: gta aaa cga cgg cca gtc.		
FEATURES			
	Location/Qualifiers		

```
source
1. .859
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNB06H07"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNB"
/Note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061
)"
BASE COUNT      238 a      172 c      199 g      250 t
ORIGIN
Query Match      17.4%; Score 686.2; DB 14; Length 859;
Best Local Similarity 88.2%; Pred. No. 1.2e-142;
Matches 758; Conservative 0; Mismatches 98; Indels 3; Gaps 1;
QY      2062 ATGCTTATGATGCCCCCAAAACAAAGAGCCACCATCAAGGACTTGCACTGCGCCCA 2121
Db      1 ATGCTTATGATGCCCCCAAAACGAAGAGCTCCGTCAGAGACTGCACTGCGCCCTA 60
QY      2122 AGTGTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2178
Db      61 AATGTGCAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY      2179 CCAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2238
Db      121 CTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY      2239 ATGCTCTTGTGGAATTGACGAGCTGCTCCAGAGCTGAGAGATGAAGAGAGAGAGAGAG 2298
Db      181 ATGCACTTGTGGAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY      2299 TAAATCAACAAAAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2358
Db      241 TAAATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY      2359 TTCTCGAGATGCTGGAACCTTGAAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTG 2418
Db      301 TTCTGAGATGCTGGAACCTTGAAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY      2419 TACATGTCATTAAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2478
Db      361 TTCAATGTCATTAAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY      2479 TCTATGATCAGTTACAGAGATATTTAAGTGTGATGAGAGAGAGAGAGAGAGAGAG 2538
Db      421 TCTACGATCAGTTACAGAGATATTTAAGTGTGATGAGAGAGAGAGAGAGAGAGAG 480
QY      2539 GGTCAATTTAGTCATACCTAAACGGGTGCTCAAGAGTTCGACCTGCAATCTTT 2598
Db      481 GGTCAATTTAGTCATACCTAAACGGGTGCTCAAGAGTTCGACCTGCAATCTTT 540
QY      2599 CAGATGCTCTCAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2658
Db      541 CTGATGCTCTCAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY      2659 ATCATGCTCTCTTGTGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 2718
Db      601 ACCATGCTCTCTTGTGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 660
QY      2719 ACATCAACTCCATCGTATCTTGGACATCTATTCCTCTTGGCTTACTGTACATTGC 2778
Db      661 ACATTAATTCATCGTATCTTGGACATCTATTCCTCTTGGCTTACTGTACATTGC 720
QY      2779 CTGCCATCTGTTTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2838
Db      721 CAGCCATCTGTTGCTAACTGGAAGTTATCACTCCAGAGCTTAAGAGAGAGAGAGAG 780
QY      2839 TGTGTTCAATGTCATTTTATCTGCAATTTTGTCTACGAGCATCTAAGAAATGAGATGA 2898
```

```
Db      781 TCTGTTTCATGTCACCTTTTATCTGATTTTGTCTACTGCGCATCTGGAATGAGATGA 840
QY      2899 GTGCTTGAATGATGA 2917
Db      841 GTGCTTGAATGATGA 859
RESULT 13
CB627056
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: L column: 20
Seq primer: gta aac cga cgg cca gtc.
FEATURES
source
1. .864
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIEB01L20"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIEB"
/Note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"
BASE COUNT      259 a      166 c      211 g      228 t
ORIGIN
Query Match      17.4%; Score 683.2; DB 14; Length 864;
Best Local Similarity 87.7%; Pred. No. 5.4e-142;
Matches 758; Conservative 0; Mismatches 103; Indels 3; Gaps 1;
QY      1744 AGAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1803
Db      1 AGAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY      1804 ATTTGTTAACTTGATTTGATCACTACATCAACACAGCAAGGCTATAAGAGAGCA 1863
Db      61 ACATGCTCACTTGATTTGATCACTACATCAACACAGCAAGGCTATAAGAGAGCA 120
QY      1864 TGTGTTTATGATGAGCCCTTACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1923
Db      121 TGTGTTTATGATGAGCCCTTGTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY      1924 GATTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1983
```



D	b		181	GGTTGATGGGATTGCATTCATGATCGTACGCCAACAGAACTGGTCTTCTTCGATA	240
O	y		1984	TCAACATGAAGAAGTTTGATGTGTAATTCAGGGTCCAAATTATGTGTA	2043
D	b		241	TCACATGAAAAGGCTTGATGTGTAATCCAAGGCCCATCTATGTGTA	300
O	y		2044	TTAGAAAGGCGAGCATTATATGTTATGATGCCCCCCAAAAACAAGACC	2103
D	b		301	TTAGAAAGGCGAGCATTATATGTTATGATGCCCCCCAAAAACAAGACC	360
O	y		2104	CTTGCAACTGCTGGCCCCAAGTGTGCTTTTGCTGTGCTTTGGCAAT	2160
D	b		361	CCTGCAACTGCTGGCCTAAATGTGCAATTTGCTGTGCTTTGGTACAG	420
O	y		2161	AAAAGAAGACTACCAAACCCAACAAGAGAAGAAAAAGTATTTTCA	2220
D	b		421	AGAAGAAACTACGAAACCTTAAACAGAGAAGAAAGAAAGATCTT	480
O	y		2221	AGAACCAATCCCCTGCATATGCTCTTGCTGAATTTGACGAAGCTG	2280
D	b		481	AAATCAATCCCCTGCATATGCACTTGCTGAATTTGAGGAGGTGCT	540
O	y		2281	ATGAAAAGCGCGGTATTGTAAATCAACAAAAATTGAAAAGAAATT	2340
D	b		541	ATGAAAAGCGCTGTATTGTAAATCAACAGAGCTGAAAAGAAATT	600
O	y		2341	TTTTTGTACATCCACACTTCTCGAAGATGTTGAACCTTGAAGT	2400
D	b		601	TTTTTGTGCATCCACACTTCTTGAAGATGTTGAACCTTGAAGT	660
O	y		2401	CTCTTTTGAAGAAGCTATACATGTCAATTAGTTGTGTTAAGA	2460
D	b		661	CTCTTTTGAAGAAGCTATACATGTCAATTAGTTGTGTTAAGA	720
O	y		2461	GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAAGATATTCT	2520
D	b		721	GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAAGATATTCT	780
O	y		2521	TGCATTGTCATGTTGGCGGTCAATTTACTGCATACCTAAACGG	2580
D	b		781	TGCATTGTCATGTTGGCGGTCAATTTACTGCATACCTAAACGG	840
O	y		2581	CTGCACCTCTGAATCTTTCAGATC	2604
D	b		841	CTGCACCACTCAATCTTTCGATC	864
<b>RESULT 14</b>					
CD437450 862 bp mRNA linear EST 03-JUN-2003					
LOCUS CD437450					
DEFINITION EL01N0501B03.b Endosperm_5 Zea mays cDNA, mRNA sequence.					
ACCESSION CD437450					
VERSION CD437450.1 GI:31353093					
KEYWORDS EST.					
SOURCE Zea mays					
ORGANISM Zea mays					
REFERENCE Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and Messing,J. Sequencing of the maize endosperm ESTS Unpublished					
TITLE JOURNAL COMMENT					
CONTACT: lai,jinsheng					
Waksman Institute, Rutgers University					
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA					
Tel: 732-445-3801					
Fax: 732-445-5735					
Email: jlai@waksman.rutgers.edu					
Seq primer: T3.					

FEATURES		Location/Qualifiers	
Source	1. 862		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/cultiyar="W2"		
	/db_xref="taxon:4577"		
	/cissue_type="Endosperm of 7-23DAP"		
	/clone_lib="Endosperm_5"		
	/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	172 a 216 c 218 g 256 t		
ORIGIN			
Query Match	17.1%; Score 672.4; DB 14; Length 862;		
Best Local Similarity	87.4%; Pred. No. 1.4e-139;		
Matches	736; Conservative 0; Mismatches 106; Indels 0; Gaps 0;		
QY	2586 CCTCTGAATCTTTCAGATCGTCTTTCACCAGGTCCTTCGGGCTCTGGGTCTATTGAG	2645	
DB	11 CCTCTCAATCTTTCGATCGTCTTTCACCAGGTCCTTCTTGGGCTCTGGTCAATTGAA	70	
QY	2646 ATCTTCTTCAGCAATCAATTGCCCTCTTGGTATGGGTATGGTGGCGGCTGAATTTTG	2705	
DB	71 ATTTCTTCAGCAACCACTGCCCTCTCTGGTATGGGTATGGTGGTGAATAAGTTCCTG	130	
QY	2706 GAAAGATTTTCCCTACATCAACTCCATCGTATCGTATCCTTATCCCTCTTGCT	2765	
DB	131 GAAAGTTTTCGTACATTAATCCATCGTATACCTTGAGACATCTATCCGCTCTTGCC	190	
QY	2766 TACTGTACATTGCCCTGCCATCTGTTTATGTACAGGAAATTTATCACTCCAGAGCTGAAT	2825	
DB	191 TATTGCACATTGCCCTGCCATCTGCTTGTGTACAGGAAATTTATCAAGCAGCTTAAC	250	
QY	2826 AATGTTGCCAGCCTGTGGTTCATGTCACTTTTATCTGCATTTTGTCTACGAGCATCTTA	2885	
DB	251 AATGTTGCCAGCCTGTGGTTCATGTCACTTTTATCTGCATTTTGTCTACGAGCATCTG	310	
QY	2886 GAAATGAGATGAGTGTGGTGTGAATTGATGACTGTGTGAAGAAAGAGAGTTCGGGTC	2945	
DB	311 GAAATGAGATGAGTGTGTGAAGCATCGATGACTGTGTGAAGAAAGAGAGTTCGGGTC	370	
QY	2946 ATTGAGGTGTGTCTCAACCTCTTGTGTGTTCAGGGACTTCTCAAGTCAATAGCT	3005	
DB	371 ATTGAGGCGGTGTCTCAACCTCTTGTGTGTTCAGGGACTTCTCAAGTCAATAGCT	430	
QY	3006 GGTTGTGATCAAGCTTCAACGTCATCATCAAAAGGTTGAGATGATGAGAGTTCAGAG	3065	
DB	431 GGTTGTGATCAAGCTTCAACGTCATCATCAAAAGGTTGAGAGATGATGAGAGTTCAGAG	490	
QY	3066 CTATATACATTCAAAATGACTACCTTATGATACCTCTTACCACTTGTCTTATGAAAC	3125	
DB	491 CTGTACACATTCAAAATGAGCACTTGTGATACCTCTTACCACTTGTCTTATGAAAC	550	
QY	3126 TTCATTGGTGTGTGCGCTGGCGTTCAAAATGCGATCAATAACGATATGAGTCATGGGGC	3185	
DB	551 TTCATTGGAGTGTGTGCGCTGGCATCTCCAATGCGATCAACAACGATATGATCATGGGGC	610	
QY	3186 CCCCTCTTTGGGAAAGCTATCTTTCATTTTGGGTGATTTGTCATCTTATCCCTTCTC	3245	
DB	611 CCCCTCTTTGGGAAAGCTATCTTTCATTTTGGGTGATTTGTCATCTTATCCCTTCTC	670	
QY	3246 AAAGTTGGTGTGAAGGCAAAACAGACACCAACGATTTGTCATGCTGTGTCATTTCTG	3305	
DB	671 AAAGTTGGTGTGAAGGCAAAACAGACACCAACGATTTGTCATGCTGTGTCATTTCTG	730	
QY	3306 CTGGCTTCAATCTTCTCGCTCTTGGGTTCGATTTGATCCTTCTTGGCAAGATGAT	3365	
DB	731 CTGGCTTCAATCTTCTCGCTCTTGGGTTCGATTTGATCCTTCTTGGCAAGATGAT	790	
QY	3366 GGTCGCTTCTTGAGAGTGTGGTGTGGATGCACTAGATGTCAAGTCAAGTCAAGTCCC	3425	
DB	791 GGTCGCTTCTTGAGAGTGTGGTGTGGATGCACTAGATGTCAAGTCAAGTCAAGTCCC	850	

QY 3426 CC 3427  
Db 851 CC 852

RESULT 15  
CC365622/c

LOCUS CC365622 900 bp DNA linear GSS 16-MAY-2003  
DEFINITION PUBR52TD ZM\_0.6\_1.0\_KB Zea mays genomic clone ZM5BTA414J08,  
genomic survey sequence.

ACCESSION CC365622  
VERSION CC365622.1 GI:30835022  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 900)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick  
,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.  
Maize Genomics Consortium  
Unpublished  
JOURNAL  
TITLE Contact: Cathy Whitelaw  
COMMENT TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.

FEATURES  
source  
1..900  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZM5BTA414J08"  
/clone\_lib="ZM\_0.6\_1.0\_KB"  
/note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
COT selected genomic DNA library"

BASE COUNT 270 a 213 c 198 g 219 t  
ORIGIN

Query Match 17.0%; Score 670.6; DB 29; Length 900;  
Best local Similarity 88.1%; Pred.No.3.6e-139;  
Matches 808; Conservative 0; Mismatches 79; Indels 30; Gaps 6;

QY 2977 TGTTCAGGAGCTTCTCAAGGTACATAGCTGTGTGATACAAAGCTTCAACGATCAA 3036  
Db 900 TGTTCAGGAGCTTCTCAAGGTACATAGCTGTGTGATACAAAGCTTCAACGATCAA 841

QY 3037 AGGTGAGATGATGAGAGTCTCAGAGCTATATACATTCAAATGAGTACCTTATTGA 3096  
Db 840 AGGTGAGATGATGAGAGTCTCAGAGCTATATACATTCAAATGAGTACCTTATTGA 781

QY 3097 TACCTCCTACCACTTGCTTCTATTGAACCTTCACTGTGTGCTGCGCTTCAATG 3156  
Db 780 TACCTCCTACCACTTGCTTCTATTGAACCTTCACTGTGTGCTGCGCTTCAATG 721

QY 3157 CGATCAATAACGATATGATGATGAGGCGCCCTCTT-TGGGAAGCTATTCTTGCATTT 3215  
Db 720 CAATCAATAATGATATGATGATGAGGCGCCCTCTT-TGGGAAGCTATTCTTGCATTT 661

QY 3216 TGGGTGATTGTCATCTTATCCCTTCTCAAGGTTTGGTGAAGCAAAACAGACA 3275  
Db 660 TGGGTGATTGTCATCTTATCCCTTCTCAAGGTTTGGTGAAGCAAAACAGACA 601

QY 3276 CCAACGATTGTCATCTGCTGCTCATTCGCTGCTTCAATCTTCTGCTTGGGTT 3335  
Db 600 CCAACGATTGTCATCTGCTGCTCATTCGCTGCTTCAATCTTCTGCTTGGGTT 541

QY 3336 CGATTGATCCTTCTCTTGGGAAGATGATGCTCCGCTTCTTGAGAGATGTGTTGAT 3395

Db 540 CGATTGACCTTCTCTTGCAAAGATGATGCTCCACTTCTTGAGAGTGTGTTGAT 481

QY 3396 TGAACCTA--GATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3453  
Db 480 TGAACCTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421

QY 3454 TCTGTGTTTGTCTCCCATATTCAGTG-----TCTGTAGATAGAGACATGAATGTC 3506  
Db 420 GCCCGCTTGTCTCCCATATTCAGTG-----TCTGTAGATAGAGACATGAATGTC 361

QY 3507 CAAGTTCTTTTGTATCCATGCTGTAACCTACTTATATCTGAGATATATCTGG-GGAAA 3565  
Db 360 CCCGTTCTTTGTATCCATGCTGTAACCTACTTATATCTGAGATATATCTGGAGGAAA 301

QY 3566 ATGAGGCTGCGGCAATCCTTGTGAGTTGGCCGTGGAATACGATATGCAAGTGT 3625  
Db 300 ATGAGGCTGCGGCAATCCTTGTGAGTTGGCCGTGGAATACGATATGCAAGTGT 241

QY 3626 GATTGTGAGCATCTTTATTTACTTGTGCGCAATATAGATGCGGTGAGCCGAAACAG 3685  
Db 240 GATTGTGAGCATCTTTATTTACTTGTGCGGCGCAAAACTGATGCTGAGCCGAGCAAG 181

QY 3686 GTATTGTGATCTGCACTGCTCCGCTGTAACAACTTGTTCTCAATAAGCAGCAGAA 3745  
Db 180 GTATTGTGATCTGCACTGCTCCGCTGTAACAACTTGTTCTCAATAAGCAGCAGAA 121

QY 3746 --TGCAATGCCAGTGGAACAGACAACTGCAATATTTATGATGCTGTTCAATTGG 3803  
Db 120 TGTCATGCTGCCAGTGGAACAGACAACTGCAATATTTATGATGCTGTTCAATTGG 75

QY 3804 AGGCTTGTTCATTTACATGTTGCTTATTAAGAAAAACAGAAATTTAGCATTAATCT 3863  
Db 74 ---CAACTTTCATTTACATGTTGCTTATTAAGAAAAACAGAAATTTAGCATTAATCT 18

QY 3864 TAGTTAATTAAGTATG 3880  
Db 17 TACTTGATTAAGTATG 1

Search completed: December 16, 2003, 00:49:16  
Job time : 7651 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 14:57:33 ; Search time 958 Seconds  
(without alignments)  
11090.802 Million cell updates/sec

Title: US-09-720-383C-9  
Perfect score: 3936  
Sequence: 1 cttctccctcgtcgtgcgg.....ctgagtcggtttgtgaaaa 3936

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 19Jun03:\*

1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3936	100.0	3936	21	AAZ58266	Corn cellulose syn
2	3936	100.0	3969	21	AAZ99503	DNA encoding a mai
3	3936	100.0	3969	21	AAZ99518	DNA encoding a mai
4	2592.6	65.9	3776	21	AAZ58263	Corn cellulose syn
5	2546.8	64.7	3568	21	AAZ99491	DNA encoding a mai
6	2241	56.9	3786	21	AAZ58265	Corn cellulose syn
7	2241	56.9	3813	21	AAZ99509	DNA encoding a mai
8	2241	56.9	3813	21	AAZ99524	DNA encoding a mai

9	1509	38.3	3255	24	ABZ14725	Arabidopsis thalia
10	1489.8	37.9	3444	22	AAC83798	Arabidopsis thalia
11	1431.8	36.4	3255	24	ABZ12754	Arabidopsis thalia
12	1429.4	36.3	3828	19	AAV06566	Arabidopsis cellul
13	1411.4	35.9	3851	21	AAA67114	Pinus radiata cell
14	1378.4	35.0	3725	21	AAZ99500	DNA encoding a mai
15	1378.4	35.0	3725	21	AAZ99506	DNA encoding a mai
16	1378.4	35.0	3725	21	AAZ99521	DNA encoding a mai
17	1364	34.7	3753	21	AAZ99515	DNA encoding a mai
18	1364	34.7	3753	21	AAZ99530	DNA encoding a mai
19	1364	34.7	3780	21	AAZ99497	DNA encoding a mai
20	1331.8	33.8	2890	21	AAZ58270	Soybean cellulose
21	1327.6	33.7	3673	19	AAV06568	Arabidopsis cellul
22	1326.6	33.7	3603	19	AAV06565	Arabidopsis cellul
23	1321.4	33.6	3704	21	AAZ99533	DNA encoding a mai
24	1289.4	32.8	3746	21	AAZ99512	DNA encoding a mai
25	1289.4	32.8	3746	21	AAZ99527	DNA encoding a mai
26	1289.4	32.8	3773	21	AAZ99494	DNA encoding a mai
27	1279.4	32.5	3799	24	AAZ58270	Corn cDNA encoding
28	1272.4	32.3	3198	21	AAC49550	Arabidopsis thalia
29	1269.2	32.2	3614	19	AAV06567	Arabidopsis cellul
30	1254.2	31.9	3517	21	AAZ58268	Soybean cellulose
31	1246.2	31.7	3311	19	AAV08372	Cellulose synthase
32	1170	29.7	2830	24	AAZ58268	Cellulose synthase
33	1137.6	28.9	3207	19	AAV08372	Cellulose synthase
34	1136.4	28.9	3328	19	AAV34432	Cotton cellulose
35	1118.8	28.4	3747	21	AAA67145	Eucalyptus grandis
36	1106.2	28.1	3232	22	AAC65448	Populus tremuloide
37	998.8	25.4	2306	19	AAV34433	Cotton cellulose
38	884.6	22.5	2248	19	AAV06562	Cellulose synthase
39	863	21.9	2033	19	AAV08381	Soybean cellulose
40	859.6	21.8	2125	21	AAZ58269	Eucalyptus grandis
41	856.4	21.8	2055	21	AAA67111	Eucalyptus grandis
42	749.2	19.0	1860	21	AAZ67081	Wheat cellulose sy
43	741.4	18.8	1029	21	AAZ58272	Wheat cellulose sy
44	701.8	17.8	1734	21	AAZ58271	Wheat cellulose sy
45	664.2	16.9	1741	19	AAV06569	Rice cellulose syn

ALIGNMENTS

RESULT 1	
AAZ58266	AAZ58266 standard; cDNA; 3936 BP.
XX	XX
AC	AAZ58266;
XX	XX
DT	08-MAY-2000 (first entry)
XX	XX
DE	Corn cellulose synthase cDNA clone p0122.ckmh70rc.
XX	XX
KW	Corn; maize; cellulose synthase; transgenic plant; ss.
XX	XX
OS	Zea mays.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	FT
XX	PN
XX	WO200004166-A2.
XX	XX
PD	27-JAN-2000.
XX	XX
PF	13-JUL-1999; 99WO-US15871.
XX	XX
PR	14-JUL-1998; 98US-0092844.
XX	XX
PA	(DUPO ) DU PONT DE NEMOURS & CO E I.
XX	XX
PI	Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G,
PI	Rafalski JA, Thorpe CJ;
XX	XX



DR MPI; 2000-182431/16.  
XX P-PSDB; AAY58835.

PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
PT probes for isolating cDNAs and genes encoding homologous proteins, for  
PT producing transgenic plants -

PS Claim 9; Page 49-50; 93pp; English.

XX The present sequence is that of full-length cDNA clone  
CC p0122.ckmh70rc encoding corn cellulose synthase (CS, see AAY58835).  
CC The cDNA clone was isolated from a cDNA library produced from corn  
CC pith tissue collected from internode subtending ear node 5 days  
CC after pollination. The invention relates to isolated nucleic acid  
CC fragments encoding plant CS and to CS polypeptides. It also relates  
CC to the construction of a chimeric gene encoding all or a portion of  
CC the CS, in sense or antisense orientation, where expression of the  
CC gene results in altered levels of the CS in transformed host cells.  
CC The host cells can be used to screen compounds for their ability to  
CC inhibit CS activity. CS nucleic acids are also useful for producing  
CC transgenic plants having altered levels of CS, and hence altered  
CC levels of fibre. CS may also serve as a target for the development  
CC of novel herbicides.

XX SQ Sequence 3936 BP; 989 A; 838 C; 1018 G; 1091 T; 0 other;

Query Match 100.0%; Score 3936; DB 21; Length 3936;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTCCCTCGTCGGTGGCGGCGGCTCGGCGTTCCGTGAGAAACCACTCGGGG 60  
DB 1 CTTCTCCCTCGTCGGTGGCGGCGGCTCGGCGTTCCGTGAGAAACCACTCGGGG 60  
QY 61 ATGAGATCTGCTGCTAGAGTGAAGAGAGCTACCGTCAAGTATCTCTGCTTCGTGGCG 120  
DB 61 ATGAGATCTGCTGCTAGAGTGAAGAGAGCTACCGTCAAGTATCTCTGCTTCGTGGCG 120  
QY 121 GCGGAAGTGGAGGGGAGAGAGCGATGAGGCGGCGGCTGTTGGCGGCTCCACA 180  
DB 121 GCGGAAGTGGAGGGGAGAGAGCGATGAGGCGGCGGCTGTTGGCGGCTCCACA 180  
QY 181 ACCGCAACGAGCTCGTCGTCATCCGCGCGAGCGGCGATCCCGGCGGAAGCGCGGG 240  
DB 181 ACCGCAACGAGCTCGTCGTCATCCGCGCGAGCGGCGATCCCGGCGGAAGCGCGGG 240  
QY 241 AGCAGAACGGGAGGTCGTCAGATTGCGGCGAGCAGTCGCGCTTGCCCGCGGGG 300  
DB 241 AGCAGAACGGGAGGTCGTCAGATTGCGGCGAGCAGTCGCGCTTGCCCGCGGGG 300  
QY 301 ACCCCTTGTGGCGGTGCAACGAGTGCGCCTTCCCGCTGCGCGGACTGCTACGAATAG 360  
DB 301 ACCCCTTGTGGCGGTGCAACGAGTGCGCCTTCCCGCTGCGCGGACTGCTACGAATAG 360  
QY 361 AGCGCCGGGAGGGCAGCAGAACTGCCCCCAGTGAAGACTGATACAAAGCGCTCAAGG 420  
DB 361 AGCGCCGGGAGGGCAGCAGAACTGCCCCCAGTGAAGACTGATACAAAGCGCTCAAGG 420  
QY 421 GCTGCCAAGTGTGACCGGTGACGAGAGAGAGAGCGGCGTGCATGACCTGGAACAAGT 480  
DB 421 GCTGCCAAGTGTGACCGGTGACGAGAGAGAGAGCGGCGTGCATGACCTGGAACAAGT 480  
QY 481 TCAACTGGAGCGGCATGACTCGAGTCTGTGCGCGAGTCCATGCTCTACGGCCCATGA 540  
DB 481 TCAACTGGAGCGGCATGACTCGAGTCTGTGCGCGAGTCCATGCTCTACGGCCCATGA 540  
QY 541 GCTACGGCCGTGAGGTGACCTTAATGGCGCGCACAAGCTTTCAGCTCAACCCCAATG 600  
DB 541 GCTACGGCCGTGAGGTGACCTTAATGGCGCGCACAAGCTTTCAGCTCAACCCCAATG 600  
QY 601 TTCCACTCTCTACCAACGGGCAATGTGTGATGACATCCACCGGAGCAGCAGCGCTGG 660  
DB 601 TTCCACTCTCTACCAACGGGCAATGTGTGATGACATCCACCGGAGCAGCAGCGCTGG 660

QY 661 TGCTTCTTTCATGGGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCA 720  
DB 661 TGCTTCTTTCATGGGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCA 720  
QY 721 GCTTACTGTGCAACCCAGGTCTATGACCATCCAAAGATCTTGTCATATGGGTATG 780  
DB 721 GCTTACTGTGCAACCCAGGTCTATGACCATCCAAAGATCTTGTCATATGGGTATG 780  
QY 781 GTAGTGTCTTGAAGAGAACGATGGAATTTGAAGCAGAGACAAGAGATGACACC 840  
DB 781 GTAGTGTCTTGAAGAGAACGATGGAATTTGAAGCAGAGACAAGAGATGACACC 840  
QY 841 AGACGGGGAATGATGTGTGTGATGATGTGACGATGCTGATCTAACCTAATGATG 900  
DB 841 AGACGGGGAATGATGTGTGTGATGATGTGACGATGCTGATCTAACCTAATGATG 900  
QY 901 AAGCAAGACAACACTGTCCAGGAAATTCACCTTCATCAAGCCAGATTAATCATATA 960  
DB 901 AAGCAAGACAACACTGTCCAGGAAATTCACCTTCATCAAGCCAGATTAATCATATA 960  
QY 961 GGATGATATCATATTATTCGCTGTGTGTTTGGGGTCTTCTTCACTACGAGTATGC 1020  
DB 961 GGATGATATCATATTATTCGCTGTGTGTTTGGGGTCTTCTTCACTACGAGTATGC 1020  
QY 1021 ATCCGGTGAATGATGATTTGCTTGTGCTCATATCTGTATCTGTAATCTGTTTG 1080  
DB 1021 ATCCGGTGAATGATGATTTGCTTGTGCTCATATCTGTATCTGTAATCTGTTTG 1080  
QY 1081 CCATGCTTGGATTCTTGATCAATTCGCAAGTGTTCCTTATGAGAGAGACTTACC 1140  
DB 1081 CCATGCTTGGATTCTTGATCAATTCGCAAGTGTTCCTTATGAGAGAGACTTACC 1140  
QY 1141 TAGACCGGCTGTCACTGAGGTTGCAAGGAAAGGCCAGCATCTCAACTGCTCCAATTG 1200  
DB 1141 TAGACCGGCTGTCACTGAGGTTGCAAGGAAAGGCCAGCATCTCAACTGCTCCAATTG 1200  
QY 1201 ATTTCTTTGTCACTACGTTGATCCCTTAAAGAACTCTTTGTGCAACAATACTTG 1260  
DB 1201 ATTTCTTTGTCACTACGTTGATCCCTTAAAGAACTCTTTGTGCAACAATACTTG 1260  
QY 1261 TTCTATCTATCCTTTCGGTGATATCTGTGATTAAGTTCTGCTATGTTCTGATG 1320  
DB 1261 TTCTATCTATCCTTTCGGTGATATCTGTGATTAAGTTCTGCTATGTTCTGATG 1320  
QY 1321 ATGTGCTGCAATGCTAAGCTTTGAAGCATTATCTGAACATCTGAATTGCAAGAAT 1380  
DB 1321 ATGTGCTGCAATGCTAAGCTTTGAAGCATTATCTGAACATCTGAATTGCAAGAAT 1380  
QY 1381 GGGTTCCTTCTGCAAAACGTAACAATTTGAACCTCGCGCTCAGAGTGTACTTCCAAC 1440  
DB 1381 GGGTTCCTTCTGCAAAACGTAACAATTTGAACCTCGCGCTCAGAGTGTACTTCCAAC 1440  
QY 1441 AGAAGATAGACTAATTGAAGACAAGGTGCGCAAACTTTGTAAGGAGAGAGAGCAA 1500  
DB 1441 AGAAGATAGACTAATTGAAGACAAGGTGCGCAAACTTTGTAAGGAGAGAGAGCAA 1500  
QY 1501 TGAAGAGAGATGAGGAATTTCAAGGTGAGATCAATGCTTACTGCAAAAGCCCAAG 1560  
DB 1501 TGAAGAGAGATGAGGAATTTCAAGGTGAGATCAATGCTTACTGCAAAAGCCCAAG 1560  
QY 1561 AAGTCTGTAAGAGAGATGAGCAATGCAATGGAACCCCTGCGCTGGAACAATGTTT 1620  
DB 1561 AAGTCTGTAAGAGAGATGAGCAATGCAATGGAACCCCTGCGCTGGAACAATGTTT 1620  
QY 1621 GTGATCATCTTGAATGATTCAGGTCTTCTTGGCCAAAGCGGAGGCTTGAATGAGG 1680  
DB 1621 GTGATCATCTTGAATGATTCAGGTCTTCTTGGCCAAAGCGGAGGCTTGAATGAGG 1680  
QY 1681 GAAATGAACCTGCAAGATTGTTTATGTTTCTAGAGAGAAACGACAGGCTATAACCATC 1740  
DB 1681 GAAATGAACCTGCAAGATTGTTTATGTTTCTAGAGAGAAACGACAGGCTATAACCATC 1740

OY 1741 ATAGAAAGCTGGTCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAATGCTC 1800  
| | | | |  
Db 1741 ATAGAAAGCTGGTCTATGAATGCATTGGTCCGAGTCTCTGCTGTACTAACAATGCTC 1800  
OY 1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATTAAGGAAG 1860  
| | | | |  
Db 1801 CATATTTGTTAAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATTAAGGAAG 1860  
OY 1861 CAATGTGTTTATGATGAGACCCCTTACTAGAAAGAGGTTTGCTATGTACAGTCCCTC 1920  
| | | | |  
Db 1861 CAATGTGTTTATGATGAGACCCCTTACTAGAAAGAGGTTTGCTATGTACAGTCCCTC 1920  
OY 1921 AAAGATTGATGGGATTGATCGCCATGACCCGATATGCTAACCCGGAATGTGCTTTTGG 1980  
| | | | |  
Db 1921 AAAGATTGATGGGATTGATCGCCATGACCCGATATGCTAACCCGGAATGTGCTTTTGG 1980  
OY 1981 ATATCAACATGAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040  
| | | | |  
Db 1981 ATATCAACATGAAGGTTTGGATGGTATTCAGGGTCCAATTTATGTTGGTACTGGATGTG 2040  
OY 2041 TATTTAGAGGCAAGCATTTATATGTTATGATGCCCCCAAAACAAGAACCCACATCAA 2100  
| | | | |  
Db 2041 TATTTAGAGGCAAGCATTTATATGTTATGATGCCCCCAAAACAAGAACCCACATCAA 2100  
OY 2101 GGACTTGCAACTGCTGCGCCCAAGTGGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGC 2160  
| | | | |  
Db 2101 GGACTTGCAACTGCTGCGCCCAAGTGGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGC 2160  
OY 2161 AAAAGAGACTACCAAAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
| | | | |  
Db 2161 AAAAGAGACTACCAAAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
OY 2221 AGAACCAATCCCTGTCATATGCTCTTGTTGAATTTGACGGAAGCTGCTCAGAGCTGAGA 2280  
| | | | |  
Db 2221 AGAACCAATCCCTGTCATATGCTCTTGTTGAATTTGACGGAAGCTGCTCAGAGCTGAGA 2280  
OY 2281 ATGAAAAGCCGGTATTGTAAATCAACAAAATTAGAAAAGAAATTGGCCAATCTTCTG 2340  
| | | | |  
Db 2281 ATGAAAAGCCGGTATTGTAAATCAACAAAATTAGAAAAGAAATTGGCCAATCTTCTG 2340  
OY 2341 TTTTGTGTACATCCACACTTCTCGAGATGTTGGAACCTTGAAGTGCAGTCCCTGCTT 2400  
| | | | |  
Db 2341 TTTTGTGTACATCCACACTTCTCGAGATGTTGGAACCTTGAAGTGCAGTCCCTGCTT 2400  
OY 2401 CTCTTTTGAAGAAGCTATACATGTCATTTAGTTGTGTTATGAACAAGACAGACTGGG 2460  
| | | | |  
Db 2401 CTCTTTTGAAGAAGCTATACATGTCATTTAGTTGTGTTATGAACAAGACAGACTGGG 2460  
OY 2461 GAAAAGAGTTGGCTGGATCTATGATCAGTTACAGAAAGATATTCTAACTGTTCAAGA 2520  
| | | | |  
Db 2461 GAAAAGAGTTGGCTGGATCTATGATCAGTTACAGAAAGATATTCTAACTGTTCAAGA 2520  
OY 2521 TGCATTTGCAATGCTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAGGTT 2580  
| | | | |  
Db 2521 TGCATTTGCAATGCTTGGCGGTCAATTTACTGCATACCTAAACGGGTTGCATTCAAGGTT 2580  
OY 2581 CTGCACCTTGAATCTTTCAAGATCGTCTTCAACCAGGTGCTTGGTGGGCTCTTGGGCTTA 2640  
| | | | |  
Db 2581 CTGCACCTTGAATCTTTCAAGATCGTCTTCAACCAGGTGCTTGGTGGGCTCTTGGGCTTA 2640  
OY 2641 TTTGAGATCTTCTTCAAGCAATCATTTGCCCTCTTTGGTATGGGTATGGGTGGGCTTGAAT 2700  
| | | | |  
Db 2641 TTTGAGATCTTCTTCAAGCAATCATTTGCCCTCTTTGGTATGGGTATGGGTGGGCTTGAAT 2700  
OY 2701 TTTTGGAAAGATTTTCTACATCAACTCATCGTGTATCTTGGACATCTATTTCCCTCT 2760  
| | | | |  
Db 2701 TTTTGGAAAGATTTTCTACATCAACTCATCGTGTATCTTGGACATCTATTTCCCTCT 2760  
OY 2761 TGGCTTACTGTACATTTGCCCTGCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGC 2820  
| | | | |  
Db 2761 TGGCTTACTGTACATTTGCCCTGCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGC 2820  
OY 2821 TGAATAATGTTGCCAGCCTGTGTTCAATGTCATCTTTTATCTGCATTTTGTGCTACGAGCA 2880  
| | | | |

Db 2821 TGAATAATGTTGCCAGCCTGTGTTCAATGTCATCTTTTATCTGCATTTTGTGCTACGAGCA 2880  
OY 2881 TCCTAGAAATGAGATGAGTGGTGTGGAATTGATGACTGCTGAGGAATGAGCAGTTCT 2940  
| | | | |  
Db 2881 TCCTAGAAATGAGATGAGTGGTGTGGAATTGATGACTGCTGAGGAATGAGCAGTTCT 2940  
OY 2941 GGGTCATTGGAGGTGTCTCTCAACCTCTTTGCTGTGTCCAGGGACTTCTCAAGTCA 3000  
| | | | |  
Db 2941 GGGTCATTGGAGGTGTCTCTCAACCTCTTTGCTGTGTCCAGGGACTTCTCAAGTCA 3000  
OY 3001 TAGCTGTGTGATACAAAGCTTCAACCGTACATCAAGGGTGAAGATGATGAGGAGTTCT 3060  
| | | | |  
Db 3001 TAGCTGTGTGATACAAAGCTTCAACCGTACATCAAGGGTGAAGATGATGAGGAGTTCT 3060  
OY 3061 CAGAGCTATATACATTCAAAATGGACTACCTTATGTATACCTTACCACCTTGCTTAT 3120  
| | | | |  
Db 3061 CAGAGCTATATACATTCAAAATGGACTACCTTATGTATACCTTACCACCTTGCTTAT 3120  
OY 3121 TGAATTCATTGGTGTGTGCTGCGCTTCAAAATGCGATCAATAACGATATGAGTCA 3180  
| | | | |  
Db 3121 TGAATTCATTGGTGTGTGCTGCGCTTCAAAATGCGATCAATAACGATATGAGTCA 3180  
OY 3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTATGTCCATCTTTATCCCT 3240  
| | | | |  
Db 3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCATTTTGGGTATGTCCATCTTTATCCCT 3240  
OY 3241 TTCTCAAAAGTTTGGTTGGAAGGCAAAACAGACACCAAGATGTCATCGTGGTCCA 3300  
| | | | |  
Db 3241 TTCTCAAAAGTTTGGTTGGAAGGCAAAACAGACACCAAGATGTCATCGTGGTCCA 3300  
OY 3301 TTCTGCTGGCTTCAATCTTCTGCTCCTTTGGGTTCCGATTGATCCTTCTTCCGAAGG 3360  
| | | | |  
Db 3301 TTCTGCTGGCTTCAATCTTCTGCTCCTTTGGGTTCCGATTGATCCTTCTTCCGAAGG 3360  
OY 3361 ATGATGCTCCGCTCTTTGAGAGAGTGTGTTGATTTGCAACTAGAGATGTCAGTGCATCAG 3420  
| | | | |  
Db 3361 ATGATGCTCCGCTCTTTGAGAGAGTGTGTTGATTTGCAACTAGAGATGTCAGTGCATCAG 3420  
OY 3421 CTCGCCCAATCTGCATATGCTTGAAGTATATTTCTGCTGTTTGTCCCATATTCAGTGT 3480  
| | | | |  
Db 3421 CTCGCCCAATCTGCATATGCTTGAAGTATATTTCTGCTGTTTGTCCCATATTCAGTGT 3480  
OY 3481 CTGTAGATTAAGAGACATGAATGTCCCAAGTTTCTTTGATCCATGTTGAACCTACTTAA 3540  
| | | | |  
Db 3481 CTGTAGATTAAGAGACATGAATGTCCCAAGTTTCTTTGATCCATGTTGAACCTACTTAA 3540  
OY 3541 TATCTGAGAGATATCTGGGGGAAATGAGGCTGCGGCAATCCTGTGCAAGTTGGGCGG 3600  
| | | | |  
Db 3541 TATCTGAGAGATATCTGGGGGAAATGAGGCTGCGGCAATCCTGTGTGCAAGTTGGGCGG 3600  
OY 3601 TGAATAACAGCATATGCAAGTGTGTTGATTTGTGAGAGCATTTCTTATTACTTGGTGGATA 3660  
| | | | |  
Db 3601 TGAATAACAGCATATGCAAGTGTGTTGATTTGTGAGAGCATTTCTTATTACTTGGTGGATA 3660  
OY 3661 TAGATGGGCTGAGCCGAAACAGAAAGTATTTGATTCTGCACTGCTCCGTTACAAACT 3720  
| | | | |  
Db 3661 TAGATGGGCTGAGCCGAAACAGAAAGTATTTGATTCTGCACTGCTCCGTTACAAACT 3720  
OY 3721 TGGTCTCAATTAAGGCAAGCAAGATGCATCTGCAAGTGAACAAGCAACCTGCACATT 3780  
| | | | |  
Db 3721 TGGTCTCAATTAAGGCAAGCAAGATGCATCTGCAAGTGAACAAGCAACCTGCACATT 3780  
OY 3781 ATTTATGTATGCTGTTCAATTGAGAGGCTTGTTCATTACATGTTGCTATACTAGAAAA 3840  
| | | | |  
Db 3781 ATTTATGTATGCTGTTCAATTGAGAGGCTTGTTCATTACATGTTGCTATACTAGAAAA 3840  
OY 3841 AACGAATATTAGCATTATCTATAGTTAATTAAAGTATGTAATGCGCTGTTTTTGT 3900  
| | | | |  
Db 3841 AACGAATATTAGCATTATCTATAGTTAATTAAAGTATGTAATGCGCTGTTTTTGT 3900  
OY 3901 TGTGTACTGTAAATCACTGAGTGTGTTTTGTGAAAA 3936  
| | | | |

Db 3901 TGTGTAATCATCTGAGTTGGTTTGTGA 3936

RESULT 2

AAZ99503

ID AAZ99503 standard; DNA; 3969 BP.

XX

AC AAZ99503;

XX

DT 03-JUL-2000 (first entry)

XX

DE DNA encoding a maize cellulose synthase.

XX

KM Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

KW transgenic plant; plant breeding marker; ss.

XX

OS Zea mays.

XX

FH Key Location/Qualifiers

FT CDS 144..3398

FT /tag= a

FT /product= "cellulose synthase"

FT /note= "no termination codon given"

XX

PN WO200009706-A2.

XX

PD 24-FEB-2000.

XX

PF 16-AUG-1999; 99WO-US18760.

XX

PR 17-AUG-1998; 98US-0096822.

XX

PA (PION-) PIONEER HI-BRED INT INC.

PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX

DR WPI; 2000-224343/19.

DR P-PSDB; AAY84111.

XX

PT New genes which encode maize cellulose synthase polypeptides in plants

PT useful for modulating the expression of cellulose synthase in plants

PT and to produce transgenic plants expressing the novel protein -

XX

PS Claim 1; Page 113-118; 119pp; English.

XX

CC The present sequence encodes a maize cellulose synthase polypeptide.

CC The cellulose synthase can be used for the improvement of stalk quality

CC for improved stand or silage. It also provides an increased concentration

CC of cellulose in the pericarp, hardening the kernel and improving its

CC handling ability. The sequences are used to produce transgenic plants

CC and seeds expressing the cellulose synthase. The polynucleotide is

CC used for modulating, preferably increasing, the level of the synthase

CC in a plant cell. The plants are preferably monocots. The polynucleotide

CC is also used as a probe or primer in the detection quantitation or

CC isolation of gene transcripts. The probes are useful in detecting

CC deficiencies in the level of mRNA in screenings for desired transgenic

CC plant, for detecting mutations in the gene, for monitoring upregulation

CC of expression or changes in enzyme activity in screening assays of

CC compounds, for detection of any number of allelic variants of the gene,

CC or for use as molecular markers in plant breeding programs. The

CC isolated nucleic acids of the present invention can also be used for

CC recombinant expression of their encoded polypeptides or for use as

CC immunogens in the preparation and/or screening of antibodies. The

CC proteins can be employed in assays for enzyme agonists or antagonists

CC of enzyme function or for use of immunogens or antigens to obtain

CC antibodies specifically immunoreactive with a protein.

XX

SQ Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 other;

Query Match 100.0%; Score 3936; DB 21; Length 3969;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTCCCTCGTCCGTGCGCGCGCTCGCGCTCGGTGAGAAACCACTCGGGG 60

Db 1 CTTCTCCCTCGTCCGTGCGCGCGCTCGCGCTCGGTGAGAAACCACTCGGGG 60

QY 61 ATGAGATCTGCTGCTAGAGTGAAGAGACTACGGTCAATATCTCTGCTTCGTGCGG 120

Db 61 ATGAGATCTGCTGCTAGAGTGAAGAGACTACGGTCAATATCTCTGCTTCGTGCGG 120

QY 121 GCGGAAGTGAAGGAGAGAAAGCGATGAGAGCGCGCGCTGCTGCGCTCCACA 180

Db 121 GCGGAAGTGAAGGAGAGAAAGCGATGAGAGCGCGCGCTGCTGCGCTCCACA 180

QY 181 ACCGCAACGAGCTCGTCTCATCCGCGGACGCGGATCCCGGCGGAGCGCGCGG 240

Db 181 ACCGCAACGAGCTCGTCTCATCCGCGGACGCGGATCCCGGCGGAGCGCGCGG 240

QY 241 AGCAGAACGGGCAAGTGTGCCAGATTGCGGCGACGACGTCGGCTTCCCCGGGGG 300

Db 241 AGCAGAACGGGCAAGTGTGCCAGATTGCGGCGACGACGTCGGCTTCCCCGGGGG 300

QY 301 ACCCTTCGTGCGCTGCAACGAGTCCGCTTCCCCGTCTGCGCGGACTGCTACGAATACG 360

Db 301 ACCCTTCGTGCGCTGCAACGAGTCCGCTTCCCCGTCTGCGCGGACTGCTACGAATACG 360

QY 361 AGCGCGGGGAGGACCGCAGAACTGCCCCAGTGCAGAACTCGATCAAGCGCCTCAAGG 420

Db 361 AGCGCGGGGAGGACCGCAGAACTGCCCCAGTGCAGAACTCGATCAAGCGCCTCAAGG 420

QY 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGAGACGCGCTCGATGACTGGAACAAGAGT 480

Db 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGAGACGCGCTCGATGACTGGAACAAGAGT 480

QY 481 TCAACTGGAGCGCCATGACTCGCAGTCTGCGCCGAGTCCATGCTCTACGGCCACATGA 540

Db 481 TCAACTGGAGCGCCATGACTCGCAGTCTGCGCCGAGTCCATGCTCTACGGCCACATGA 540

QY 541 GCTACGCGCGTGAAGGTGACCCCTAATGCGCGCCACCAAGCTTTCAGCTCAACCCCAATG 600

Db 541 GCTACGCGCGTGAAGGTGACCCCTAATGCGCGCCACCAAGCTTTCAGCTCAACCCCAATG 600

QY 601 TTCCACTCTCAACCAACGCGCAATGTGTGATGACATCCCAACGAGAGACGCGCTGG 660

Db 601 TTCCACTCTCAACCAACGCGCAATGTGTGATGACATCCCAACGAGAGACGCGCTGG 660

QY 661 TGCCTTCTTCAATGGGTGTGGGGAAGAGATACATCCCTTCTTATGCGATCCCA 720

Db 661 TGCCTTCTTCAATGGGTGTGGGGAAGAGATACATCCCTTCTTATGCGATCCCA 720

QY 721 GCTTACCTGTGCAACCGGTCTATGACCCATCCAGGATCTTGCTGATATGGGTATG 780

Db 721 GCTTACCTGTGCAACCGGTCTATGACCCATCCAGGATCTTGCTGATATGGGTATG 780

QY 781 GTAGTGTGCTTGAAGAGAACGGATGAGAAATTGGAAGCAGAGACAGAGAGATGCACC 840

Db 781 GTAGTGTGCTTGAAGAGAACGGATGAGAAATTGGAAGCAGAGACAGAGAGATGCACC 840

QY 841 AGACGGGGAATGATGTGTGTGTGATGATGTGACGATGCTGATCTACCACTAATGATG 900

Db 841 AGACGGGGAATGATGTGTGTGTGATGATGTGACGATGCTGATCTACCACTAATGATG 900

QY 901 AAGCAAGACAACAACCTGTCAGGAAATTCACCTTCATCAAGCCAGATTAATCCATATA 960

Db 901 AAGCAAGACAACAACCTGTCAGGAAATTCACCTTCATCAAGCCAGATTAATCCATATA 960

QY 961 GGATGATTATCATTAATTCGGCTTGTGTTGGGGTCTTCTTCCACTACGAGTGATGC 1020

Db 961 GGATGATTATCATTAATTCGGCTTGTGTTGGGGTCTTCTTCCACTACGAGTGATGC 1020

QY 1021 ATCCGCTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGTGAATCTGTTTG 1080

Db 1021 ATCCGCTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGTGAATCTGTTTG 1080

QY 1081 CCATGTCTTGATTTCTTGATCAATTCCAAAGTGTTCCCTATTGAGAGAGAGACTTACC 1140



Db 1081 ||||| CCATGCTTGGATTCTTGATCAATTCCAAAGTGTTCCCTATTGAGAGAGACTTACC 1140  
Qy 1141 TAGACCGCGTGTCACTGAGGTTGACAGAGAAAGCCAGCCATCTCAACTGTCTCAATTG 1200  
|||  
Db 1141 TAGACCGCGTGTCACTGAGGTTGACAGAGAAAGCCAGCCATCTCAACTGTCTCAATTG 1200  
Qy 1201 ATTTCTTGTGAGTACGGTTGATCCCTTAAGAACCTCTTGTGTCAACAATAACTG 1260  
|||  
Db 1201 ATTTCTTGTGAGTACGGTTGATCCCTTAAGAACCTCTTGTGTCAACAATAACTG 1260  
Qy 1261 TTCTATCTATCCTTTCCGGTGATTAATCCTGTGTGATAAGGTTCTTGTCTATGTTCTGATG 1320  
|||  
Db 1261 TTCTATCTATCCTTTCCGGTGATTAATCCTGTGTGATAAGGTTCTTGTCTATGTTCTGATG 1320  
Qy 1321 ATGGTGTGCAATGCTTAACGTTTGAAGCAATTATCTGAACATCTGAATTGTCCAAAGAAAT 1380  
|||  
Db 1321 ATGGTGTGCAATGCTTAACGTTTGAAGCAATTATCTGAACATCTGAATTGTCCAAAGAAAT 1380  
Qy 1381 GGGTTCCTTTCTGCAAAACGGTACAAATATGAACTCTCGCGCTCCAGAGTGTACTTCCAAC 1440  
|||  
Db 1381 GGGTTCCTTTCTGCAAAACGGTACAAATATGAACTCTCGCGCTCCAGAGTGTACTTCCAAC 1440  
Qy 1441 AGAAGATGACTACTTGAAGAACAAGTCCAGCAAACTTTGTTAGGAGAGAGAGACAA 1500  
|||  
Db 1441 AGAAGATGACTACTTGAAGAACAAGTCCAGCAAACTTTGTTAGGAGAGAGAGACAA 1500  
Qy 1501 TGAAGAGAGATATGAGGAATTCAGGTGACAAATCAATGCCCTTAGTTGCCAAAGCCAG 1560  
|||  
Db 1501 TGAAGAGAGATATGAGGAATTCAGGTGACAAATCAATGCCCTTAGTTGCCAAAGCCAG 1560  
Qy 1561 AAGTTCCTGAAGAAGATGAGCAATGCAAGATGGAACCCCTGGCTGGAAACAATGTTT 1620  
|||  
Db 1561 AAGTTCCTGAAGAAGATGAGCAATGCAAGATGGAACCCCTGGCTGGAAACAATGTTT 1620  
Qy 1621 GTGATCATCTGGAATGATTCAAGTCTTCTGGCCAAAGCGAGCCTTGACTGTGAGG 1680  
|||  
Db 1621 GTGATCATCTGGAATGATTCAAGTCTTCTGGCCAAAGCGAGCCTTGACTGTGAGG 1680  
Qy 1681 GAAATGAACCTGCCAGATTGTTTATGTTCTAGAGAGAAACGACGACTATAACCATC 1740  
|||  
Db 1681 GAAATGAACCTGCCAGATTGTTTATGTTCTAGAGAGAAACGACGACTATAACCATC 1740  
Qy 1741 ATAAGAAAGCTGCTGCTATGAATGCATTGCTCCGAGTCTCTGCTGTACTAACAATGCTC 1800  
|||  
Db 1741 ATAAGAAAGCTGCTGCTATGAATGCATTGCTCCGAGTCTCTGCTGTACTAACAATGCTC 1800  
Qy 1801 CATATTTGTTAAACCTTGAGTTGTGATCACTAACAACAAGCAGCAAGCTATAAGGAAG 1860  
|||  
Db 1801 CATATTTGTTAAACCTTGAGTTGTGATCACTAACAACAAGCAGCAAGCTATAAGGAAG 1860  
Qy 1861 CAATGTGTTTATGATGAGACCCCTTACTAGAGAAAGAGTTGTGTAAGATTCCCTC 1920  
|||  
Db 1861 CAATGTGTTTATGATGAGACCCCTTACTAGAGAAAGAGTTGTGTAAGATTCCCTC 1920  
Qy 1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTTTTTG 1980  
|||  
Db 1921 AAAGATTTGATGGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGTTTTTG 1980  
Qy 1981 ATATCAACATGAAAGGTTTGATGGTATTCAGGGTCCAAATTTATGTTGTAAGTGTG 2040  
|||  
Db 1981 ATATCAACATGAAAGGTTTGATGGTATTCAGGGTCCAAATTTATGTTGTAAGTGTG 2040  
Qy 2041 TATTTAGAAAGCAGGCAATTATATGTTATGATGCCCCCAAAACAAGAGCCACCATCAA 2100  
|||  
Db 2041 TATTTAGAAAGCAGGCAATTATATGTTATGATGCCCCCAAAACAAGAGCCACCATCAA 2100  
Qy 2101 GGACTTGCAACTGCTGGCCCAAGTGTTGCTGTGCTGTGCTTGGCAATAGGAAGC 2160  
|||  
Db 2101 GGACTTGCAACTGCTGGCCCAAGTGTTGCTGTGCTGTGCTTGGCAATAGGAAGC 2160  
Qy 2161 AAAAGAGACTACCAAAACCAAGAGAGAGAAAGTATATTTTCAAGAAAGAG 2220  
|||

Db 2161 AAAAGAGACTACCAAAACCAAGAGAGAGAGAAAGTATATTTTCAAGAAAGAG 2220  
Qy 2221 AGAACCAATCCCCCTGCATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGAGCTGAG 2280  
|||  
Db 2221 AGAACCAATCCCCCTGCATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGAGCTGAG 2280  
Qy 2281 ATGAAAAGCCGGTATTGTAAATCAACAAAATTAGAAAAGAAATTGGCCAATCTTCTG 2340  
|||  
Db 2281 ATGAAAAGCCGGTATTGTAAATCAACAAAATTAGAAAAGAAATTGGCCAATCTTCTG 2340  
Qy 2341 TTTTGTTCATCCACACTTCTCGAAGATGTGGAACCTTGAAGAGTGCAAGTCTGCTT 2400  
|||  
Db 2341 TTTTGTTCATCCACACTTCTCGAAGATGTGGAACCTTGAAGAGTGCAAGTCTGCTT 2400  
Qy 2401 CTCTTTGAAAGAGCTATACATGTCTATTAGTTGTGTTAGAAACAAGACAGACTGGG 2460  
|||  
Db 2401 CTCTTTGAAAGAGCTATACATGTCTATTAGTTGTGTTAGAAACAAGACAGACTGGG 2460  
Qy 2461 GAAAGAGATTGGCTGGATCTATGATCAGTTACAGAGATATTTCTAAGTGTTCAGAG 2520  
|||  
Db 2461 GAAAGAGATTGGCTGGATCTATGATCAGTTACAGAGATATTTCTAAGTGTTCAGAG 2520  
Qy 2521 TGCATTGTCAATGTTGGCGGTCAATTTACTGCATACCTTAAACGGGTGCATTCAAGGTT 2580  
|||  
Db 2521 TGCATTGTCAATGTTGGCGGTCAATTTACTGCATACCTTAAACGGGTGCATTCAAGGTT 2580  
Qy 2581 CTGCACTCTGAATCTTTCAGATCGTCTTCAACAGGTCCTCGTGCGCTCTTGGGCTTA 2640  
|||  
Db 2581 CTGCACTCTGAATCTTTCAGATCGTCTTCAACAGGTCCTCGTGCGCTCTTGGGCTTA 2640  
Qy 2641 TTGAGATCTTCTTCAAGATCAATGCCCTCTTGTGATGGGTATGGTGCGGTGAAAT 2700  
|||  
Db 2641 TTGAGATCTTCTTCAAGATCAATGCCCTCTTGTGATGGGTATGGTGCGGTGAAAT 2700  
Qy 2701 TTTTGAAGAAGATTTTCTTACATCAACTCCATCGTGTATCTTGGACATCTATTTCCCTCT 2760  
|||  
Db 2701 TTTTGAAGAAGATTTTCTTACATCAACTCCATCGTGTATCTTGGACATCTATTTCCCTCT 2760  
Qy 2761 TGGCTTACTGTACATTTGCCCTGCACTCTGTTATTATGACAGGAAATTTATCACTCCAGAGC 2820  
|||  
Db 2761 TGGCTTACTGTACATTTGCCCTGCACTCTGTTATTATGACAGGAAATTTATCACTCCAGAGC 2820  
Qy 2821 TGAATATGTTGCCAGCCTGTGGTTCAATGTCACTTTTATCTGCAATTTTGTCTACGAGCA 2880  
|||  
Db 2821 TGAATATGTTGCCAGCCTGTGGTTCAATGTCACTTTTATCTGCAATTTTGTCTACGAGCA 2880  
Qy 2881 TCCTAGAAATGAGATGAGTGGTGTGGAATTGATGACTGGTGAGAGAAATGAGCAAGTTCT 2940  
|||  
Db 2881 TCCTAGAAATGAGATGAGTGGTGTGGAATTGATGACTGGTGAGAGAAATGAGCAAGTTCT 2940  
Qy 2941 GGGTCAATTGGAGGTGTGTCTCACACCTCTTGTGTGTGTCCAGGGAAGTTCTCAAGGTCA 3000  
|||  
Db 2941 GGGTCAATTGGAGGTGTGTCTCACACCTCTTGTGTGTGTCCAGGGAAGTTCTCAAGGTCA 3000  
Qy 3001 TAGCTGGTGTGATACAAAGCTTCAACGTGACATCAAAAGGTGAGATGAGAGAGTTCT 3060  
|||  
Db 3001 TAGCTGGTGTGATACAAAGCTTCAACGTGACATCAAAAGGTGAGATGAGAGAGTTCT 3060  
Qy 3061 CAGAGCTATATACATTCAATGACTACCTTATTTGATACCTCTTACCACTTGTCTTAT 3120  
|||  
Db 3061 CAGAGCTATATACATTCAATGACTACCTTATTTGATACCTCTTACCACTTGTCTTAT 3120  
Qy 3121 TGAATTCATTTGTTGTTGCTGGCTGCTTCAAAATGCGATCAATAACGATATGAGTCA 3180  
|||  
Db 3121 TGAATTCATTTGTTGTTGCTGGCTGCTTCAAAATGCGATCAATAACGATATGAGTCA 3180  
Qy 3181 GGGGCCCTCTTTGGGAAGCTATTCTTGCATTTTGGGTATGTCATCTTATCCCT 3240  
|||  
Db 3181 GGGGCCCTCTTTGGGAAGCTATTCTTGCATTTTGGGTATGTCATCTTATCCCT 3240  
Qy 3241 TTCTCAAAAGTTTGTGGAAGGCAAAACAGACACCAACGATGTCATGCTGTGTC 3300  
|||  
Db 3241 TTCTCAAAAGTTTGTGGAAGGCAAAACAGACACCAACGATGTCATGCTGTGTC 3300

QY 3301 TTCTGCTGGCTTCAATCTTCTCGCTCCCTTGGGTTGGATTGATCCCTTTCCTTGGCAAG 3360  
Db 3301 TTCTGCTGGCTTCAATCTTCTCGCTCCCTTGGGTTGGATTGATCCCTTTCCTTGGCAAG 3360  
QY 3361 ATGATGTCGCTTCTTGAGAGTGTGGTTGGATTGCAACTAGAGTTCAGTGCATCAG 3420  
Db 3361 ATGATGTCGCTTCTTGAGAGTGTGGTTGGATTGCAACTAGAGTTCAGTGCATCAG 3420  
QY 3421 CTCGCCCAATCGCATATGCTTGAAGTATTTCTGGTGTGGTTGGCCCATATTCAAGTGT 3480  
Db 3421 CTCGCCCAATCGCATATGCTTGAAGTATTTCTGGTGTGGTTGGCCCATATTCAAGTGT 3480  
QY 3481 CTGTAGATAAGACATGAATGTCCCAAGTTCTTTTGATCCATGGTGAACTACTTAA 3540  
Db 3481 CTGTAGATAAGACATGAATGTCCCAAGTTCTTTTGATCCATGGTGAACTACTTAA 3540  
QY 3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGGGCAATCCTTGTGAGTTGGCCG 3600  
Db 3541 TATCTGAGAGATATACTGGGGGAAAATGGAGGCTGGGCAATCCTTGTGAGTTGGCCG 3600  
QY 3601 TGAATACAGCATATGCAAGTGTGATTGTGAGCATTTCTTATTACTTGGTGCATA 3660  
Db 3601 TGAATACAGCATATGCAAGTGTGATTGTGAGCATTTCTTATTACTTGGTGCATA 3660  
QY 3661 TAGATGGCTGAGCCGAACAGCAAGGATTTGATTTCTGCACTGCTCCGTTGACAACT 3720  
Db 3661 TAGATGGCTGAGCCGAACAGCAAGGATTTGATTTCTGCACTGCTCCGTTGACAACT 3720  
QY 3721 TGGTCTCAATAAGCAGCAGGAGATGCATCTGCCAGTGAACAGAGCAACTGCACATT 3780  
Db 3721 TGGTCTCAATAAGCAGCAGGAGATGCATCTGCCAGTGAACAGAGCAACTGCACATT 3780  
QY 3781 ATTTATGATGCTGTTCATTGGAGGGCTGTTCATTACATGTCTGCTATACTAGAAAA 3840  
Db 3781 ATTTATGATGCTGTTCATTGGAGGGCTGTTCATTACATGTCTGCTATACTAGAAAA 3840  
QY 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAGTATGTAATGGCGCTGTTTGT 3900  
Db 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAGTATGTAATGGCGCTGTTTGT 3900  
QY 3901 TGTGACTGTAATCATCTGAGTTGTTTGTGAAA 3936  
Db 3901 TGTGACTGTAATCATCTGAGTTGTTTGTGAAA 3936

RESULT 3  
AAZ99518  
ID AAZ99518 standard; DNA; 3969 BP.  
XX  
AC AAZ99518;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE DNA encoding a maize cellulose synthase.  
XX  
KM Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker; ss.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 144..3398  
FT /\*tag= a /product= "cellulose synthase"  
FT /note= "no termination codon given"  
XX  
PN WO200009706-A2.  
XX  
PD 24-FEB-2000.  
XX  
PF 16-AUG-1999; 99WO-US18760.  
XX

PR 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Dhurga KS, Helentjaris TG, Bowen BA, Wang X;  
XX  
DR WPI; 2000-224343/19.  
DR P-PSDB; AAY84116.  
XX  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein  
XX  
PS Claim 1; Page 152-157; 119pp; English.  
XX  
CC The present sequence encodes a maize cellulose synthase polypeptide.  
CC The cellulose synthase can be used for the improvement of stalk quality  
CC for improved stand or silage. It also provides an increased concentration  
CC of cellulose in the pericarp, hardening the kernel and improving its  
CC handling ability. The sequences are used to produce transgenic plants  
CC and seeds expressing the cellulose synthase. The polynucleotide is  
CC used for modulating, preferably increasing, the level of the synthase  
CC in a plant cell. The plants are preferably monocots. The polynucleotide  
CC is also used as a probe or primer in the detection quantitation or  
CC isolation of gene transcripts. The probes are useful in detecting  
CC deficiencies in the level of mRNA in screenings for desired transgenic  
CC plant, for detecting mutations in the gene, for monitoring upregulation  
CC of expression or changes in enzyme activity in screening assays of  
CC compounds, for detection of any number of allelic variants of the gene,  
CC or for use as molecular markers in plant breeding programs. The  
CC isolated nucleic acids of the present invention can also be used for  
CC recombinant expression of their encoded polypeptides or for use as  
CC immunogens in the preparation and/or screening of antibodies. The  
CC proteins can be employed in assays for enzyme agonists or antagonists  
CC of enzyme function or for use of immunogens or antigens to obtain  
CC antibodies specifically immunoreactive with a protein.  
XX  
SQ Sequence 3969 BP; 1022 A; 838 C; 1018 G; 1091 T; 0 other;  
Query Match 100.0%; Score 3936; DB 21; Length 3969;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCTCCCTGCTGCTGAGTGAAGAGCTACGCTCAGTATCCTTGCTTCTGCGCG 60  
Db 1 CTTCTCCCTGCTGCTGAGTGAAGAGCTACGCTCAGTATCCTTGCTTCTGCGCG 60  
QY 61 ATGAGATCTGCTGCTAGAGTGAAGAGCTACGCTCAGTATCCTTGCTTCTGCGCG 120  
Db 61 ATGAGATCTGCTGCTAGAGTGAAGAGCTACGCTCAGTATCCTTGCTTCTGCGCG 120  
QY 121 GCGGAGTGAAGGAGGAGGAGCGATGAGCGCGCGGCTGTGCGCGCTCCACA 180  
Db 121 GCGGAGTGAAGGAGGAGGAGCGATGAGCGCGCGGCTGTGCGCGCTCCACA 180  
QY 181 ACCGCAAGAGCTGCTGTCATCCGCGCGAGCGGCGATCCCGGCGGAAGCGCGCGG 240  
Db 181 ACCGCAAGAGCTGCTGTCATCCGCGCGAGCGGCGATCCCGGCGGAAGCGCGCGG 240  
QY 241 AGCAGAACGCGCAGGTGTGCCAGATTGCGGCGAGCAGCTCGGCTTGCCCCGCGGG 300  
Db 241 AGCAGAACGCGCAGGTGTGCCAGATTGCGGCGAGCAGCTCGGCTTGCCCCGCGGG 300  
QY 301 ACCCTTCTGCTGCGTGAACGAGTGGCCTTCCCCGCTTGCGCGGAGCTGCAAGATACG 360  
Db 301 ACCCTTCTGCTGCGTGAACGAGTGGCCTTCCCCGCTTGCGCGGAGCTGCAAGATACG 360  
QY 361 AGCGCGGAGGCGACGAGAACTGCCCCAGTGCAGAGCTGCATACAAGCGCTCAAGG 420  
Db 361 AGCGCGGAGGCGACGAGAACTGCCCCAGTGCAGAGCTGCATACAAGCGCTCAAGG 420  
QY 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGGAGCGGCGTGCATGACCTGACAAACGAGT 480  
Db 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGGAGCGGCGTGCATGACCTGACAAACGAGT 480

Db 421 GCTGCCAACGTGTGACCCGGTGACGAGGAGGACGGCGTGTGATGACCTTGACAACGAGT 480  
Qy 481 TCAACTGGGACGGCCATGACTCGAGTCTGTGGCCGAGTGCATGCTTACGGCCACATGA 540  
Db 481 TCAACTGGGACGGCCATGACTCGAGTCTGTGGCCGAGTGCATGCTTACGGCCACATGA 540  
Qy 541 GCTACGGCCGTGGAGGTGACCCCTAATGGCGGCCACAAGCTTTCAGCTCAACCCCAATG 600  
Db 541 GCTACGGCCGTGGAGGTGACCCCTAATGGCGGCCACAAGCTTTCAGCTCAACCCCAATG 600  
Qy 601 TTCCACTCCTCACCAACGGGCAAAATGATGATGACATCCCAACCGAGCAGCAGCGCGTGG 660  
Db 601 TTCCACTCCTCACCAACGGGCAAAATGATGATGACATCCCAACCGAGCAGCAGCGCGTGG 660  
Qy 661 TGCCTTCTTTCATGGGTGGGTGGGGAAGAAGATACATCCCTTTCCTTATGCGGATCCCA 720  
Db 661 TGCCTTCTTTCATGGGTGGGTGGGGAAGAAGATACATCCCTTTCCTTATGCGGATCCCA 720  
Qy 721 GCTTACCTGTGCAACCCAGGTCTATGGAACCCATCCAAAGGATCTGTGTCATATGGGTATG 780  
Db 721 GCTTACCTGTGCAACCCAGGTCTATGGAACCCATCCAAAGGATCTGTGTCATATGGGTATG 780  
Qy 781 GTAGTGTCTTGTGAAGGAACGATGGAATTTGAAGCAGAGACAAGAGAGATGCACC 840  
Db 781 GTAGTGTCTTGTGAAGGAACGATGGAATTTGAAGCAGAGACAAGAGAGATGCACC 840  
Qy 841 AGACGGGGAATGATGGTGGTGGTGAATGATGGTGAACGATGCTGATCTAACCTAATGATG 900  
Db 841 AGACGGGGAATGATGGTGGTGGTGAATGATGGTGAACGATGCTGATCTAACCTAATGATG 900  
Qy 901 AAGCAGACAAACAACCTGTCAGAAAAATTCACCTTCATCAAGCCAGATTAAATCCATATA 960  
Db 901 AAGCAGACAAACAACCTGTCAGAAAAATTCACCTTCATCAAGCCAGATTAAATCCATATA 960  
Qy 961 GGATGATTAATCATTATTCCGCTTGTGGTTTGGGGTCTTCTTCCACTACCGAGTGAATG 1020  
Db 961 GGATGATTAATCATTATTCCGCTTGTGGTTTGGGGTCTTCTTCCACTACCGAGTGAATG 1020  
Qy 1021 ATCCGGTGAATGATGCAATTTGTGTGTGGCTCATATCTGTTAATCTGTGAATCTGGTTTG 1080  
Db 1021 ATCCGGTGAATGATGCAATTTGTGTGTGGCTCATATCTGTTAATCTGTGAATCTGGTTTG 1080  
Qy 1081 CCATGTCCTTGGATTCTTGATCAATTCGCAAAAGTGTTCCCTAATGAGAGAGAGACTTACC 1140  
Db 1081 CCATGTCCTTGGATTCTTGATCAATTCGCAAAAGTGTTCCCTAATGAGAGAGAGACTTACC 1140  
Qy 1141 TAGACCGGCTGTCACTGAGTTTCGACAAAGGAGCCAGCCATCTCAACTGTCTCCAATTG 1200  
Db 1141 TAGACCGGCTGTCACTGAGTTTCGACAAAGGAGCCAGCCATCTCAACTGTCTCCAATTG 1200  
Qy 1201 ATTTCTTTGTCACTACGTTGATCCCTTAAAGGAACCTCCTTGTGTCACAACAATACTAG 1260  
Db 1201 ATTTCTTTGTCACTACGTTGATCCCTTAAAGGAACCTCCTTGTGTCACAACAATACTAG 1260  
Qy 1261 TTCTAATCTATCCTTTCGGTGGATTAATCCTGTGTGATAAGGTTTCTGTATGTCTGATG 1320  
Db 1261 TTCTAATCTATCCTTTCGGTGGATTAATCCTGTGTGATAAGGTTTCTGTATGTCTGATG 1320  
Qy 1321 ATGTTGCTGCAATGCTAACGTTTGAAGCAATTAATCTGAACAATCTGAATTTGCCAAGAAAT 1380  
Db 1321 ATGTTGCTGCAATGCTAACGTTTGAAGCAATTAATCTGAACAATCTGAATTTGCCAAGAAAT 1380  
Qy 1381 GGGTTCCTTTCGCAAAACGGTACAAATATGAAACCTCGCGCTCAGAGTGTACTTCCAAC 1440  
Db 1381 GGGTTCCTTTCGCAAAACGGTACAAATATGAAACCTCGCGCTCAGAGTGTACTTCCAAC 1440  
Qy 1441 AGAAGATAGACTACTTGAAGAACAAGTGGCAGCAAACTTTGTTAGGGAAGAGAGAGACAA 1500  
Db 1441 AGAAGATAGACTACTTGAAGAACAAGTGGCAGCAAACTTTGTTAGGGAAGAGAGACAA 1500  
Qy 1501 TGAAGAGAGATATGAGGAATTCAGGTGAGAAATCAATGCTTATGTTGCCAAAGCCCAAGA 1560  
Db 1501 TGAAGAGAGATATGAGGAATTCAGGTGAGAAATCAATGCTTATGTTGCCAAAGCCCAAGA 1560

Qy 1561 AAGTTCCTGAAGAAGATGGAACAATGCAAGATGMAACCCCTGGCCTGGAAACAATGTTTC 1620  
Db 1561 AAGTTCCTGAAGAAGATGGAACAATGCAAGATGMAACCCCTGGCCTGGAAACAATGTTTC 1620  
Qy 1621 GTGATCATCCTGGAATGATTCAAGTCTTCTTGGGCCAAAGCGGAGGCTTGAATGTGAGG 1680  
Db 1621 GTGATCATCCTGGAATGATTCAAGTCTTCTTGGGCCAAAGCGGAGGCTTGAATGTGAGG 1680  
Qy 1681 GAAATGAACCTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACAGGCTATTAACCATC 1740  
Db 1681 GAAATGAACCTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACAGGCTATTAACCATC 1740  
Qy 1741 ATAGAAAAGCTGGTCTAATGAATGCATTGGTCCGAGTCTCTGCTGTACTTAACAATGCTC 1800  
Db 1741 ATAGAAAAGCTGGTCTAATGAATGCATTGGTCCGAGTCTCTGCTGTACTTAACAATGCTC 1800  
Qy 1801 CATATTTGTTAACTTGATTTGTGATCACTACATCAACAACGACAGGCTATTAAGGAAG 1860  
Db 1801 CATATTTGTTAACTTGATTTGTGATCACTACATCAACAACGACAGGCTATTAAGGAAG 1860  
Qy 1861 CAATGTCTTTATGATGGAACCCCTTACTAGGAAGAAGGTTTGTATGTACAGTTCCTTC 1920  
Db 1861 CAATGTCTTTATGATGGAACCCCTTACTAGGAAGAAGGTTTGTATGTACAGTTCCTTC 1920  
Qy 1921 AAAGATTTGATGGATGATGATGCCATGACCCGATATGCTAAACCGAATGTGTCTTTTGG 1980  
Db 1921 AAAGATTTGATGGATGATGATGCCATGACCCGATATGCTAAACCGAATGTGTCTTTTGG 1980  
Qy 1981 ATATCAACATGAAAAGTTTGGATGGTATTCAAGGTCCAAATTTATGTTGGTACTGATGTG 2040  
Db 1981 ATATCAACATGAAAAGTTTGGATGGTATTCAAGGTCCAAATTTATGTTGGTACTGATGTG 2040  
Qy 2041 TATTTAGAAGGAGGCAATTATATGTTATGATGCCCCCAAAAACAAGAGCCCATCA 2100  
Db 2041 TATTTAGAAGGAGGCAATTATATGTTATGATGCCCCCAAAAACAAGAGCCCATCA 2100  
Qy 2101 GGACTTGCACACTGTGCGCCCAAGTGGTCTTTTGCTGTGCTGCTTGGCAATAGGAAGC 2160  
Db 2101 GGACTTGCACACTGTGCGCCCAAGTGGTCTTTTGCTGTGCTGCTTGGCAATAGGAAGC 2160  
Qy 2161 AAAAGAAGACTACCAAAACCCAAAACAGAGAAGAAAAGTTATTTTCAAGAAAGAG 2220  
Db 2161 AAAAGAAGACTACCAAAACCCAAAACAGAGAAGAAAAGTTATTTTCAAGAAAGAG 2220  
Qy 2221 AGAACCAATCCCTGCAATATGCTCTTGTAATGACGAAGCTGCTCCAGAGCTGAGA 2280  
Db 2221 AGAACCAATCCCTGCAATATGCTCTTGTAATGACGAAGCTGCTCCAGAGCTGAGA 2280  
Qy 2281 ATGAAAAAGCCGGTATTTGTAATCAACAAAAATTAGAAAAAGAAATTGGCCAAATCTTCTG 2340  
Db 2281 ATGAAAAAGCCGGTATTTGTAATCAACAAAAATTAGAAAAAGAAATTGGCCAAATCTTCTG 2340  
Qy 2341 TTTTGTACATCCACACTTCTCGAAGATGGTGAACCTGAAGAGTGCAAGTCTGCTT 2400  
Db 2341 TTTTGTACATCCACACTTCTCGAAGATGGTGAACCTGAAGAGTGCAAGTCTGCTT 2400  
Qy 2401 CTCTTTGAAAGAAGCTATACATGTCAATTAATGTTGTGTTATGAAGACAAGACAGACTGGG 2460  
Db 2401 CTCTTTGAAAGAAGCTATACATGTCAATTAATGTTGTGTTATGAAGACAAGACAGACTGGG 2460  
Qy 2461 GAAAAAGAGATTGCGTGAATCTATGATCAGTTACAGAGATATTCTAATGTTTGAAGA 2520  
Db 2461 GAAAAAGAGATTGCGTGAATCTATGATCAGTTACAGAGATATTCTAATGTTTGAAGA 2520  
Qy 2521 TGCATTTGTCATGTTGGCGGTCAATTTACTGTCAATCTAAACGGGTTGCATTCAAAGGTT 2580  
Db 2521 TGCATTTGTCATGTTGGCGGTCAATTTACTGTCAATCTAAACGGGTTGCATTCAAAGGTT 2580  
Qy 2581 CTGCACTCTGAATCTTTCAGATCGTCTTCAACCAAGTGTCTCGGTGGGCTCTTGGGCTTA 2640  
Db 2581 CTGCACTCTGAATCTTTCAGATCGTCTTCAACCAAGTGTCTCGGTGGGCTCTTGGGCTTA 2640



```

QY 2641 TTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGGTGGCGCTGTAAT 2700
Db 2641 TTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGGTGGCGCTGTAAT 2700
QY 2701 TTTTGAAGAATTTTCTACATCACTCCATCGTGTATCCTTGACACTATATCCCTCT 2760
Db 2701 TTTTGAAGAATTTTCTACATCACTCCATCGTGTATCCTTGACACTATATCCCTCT 2760
QY 2761 TGGCTTACTGTACATTCGCTGCCATCTGTTATTTGACAGGAAATTTATCTCCAGAGC 2820
Db 2761 TGGCTTACTGTACATTCGCTGCCATCTGTTATTTGACAGGAAATTTATCTCCAGAGC 2820
QY 2821 TGAATATGTTCGACGCCCTGTGTTCACTGCTATTTATCTGCAATTTTGTCTACGAGCA 2880
Db 2821 TGAATATGTTCGACGCCCTGTGTTCACTGCTATTTATCTGCAATTTTGTCTACGAGCA 2880
QY 2881 TCCTAGAAATGAGATGAGTGTGTTGGAATGATGACTGTGAGGAAATGAGCAAGTTCT 2940
Db 2881 TCCTAGAAATGAGATGAGTGTGTTGGAATGATGACTGTGAGGAAATGAGCAAGTTCT 2940
QY 2941 GGGTCATGAGGTGTGCTCTCACACCTCTTGTCTGTCTCCAGGACTTCTCAAGTCA 3000
Db 2941 GGGTCATGAGGTGTGCTCTCACACCTCTTGTCTGTCTCCAGGACTTCTCAAGTCA 3000
QY 3001 TAGCTGTGTGATACAGCTTCAACCGTGACATCAAGGGTGAGATGATGAGGAGTTCT 3060
Db 3001 TAGCTGTGTGATACAGCTTCAACCGTGACATCAAGGGTGAGATGATGAGGAGTTCT 3060
QY 3061 CAGAGCTATATACATTCAAATGAGACTACCTTATGATACCTCTACACCTTGTCTTAT 3120
Db 3061 CAGAGCTATATACATTCAAATGAGACTACCTTATGATACCTCTACACCTTGTCTTAT 3120
QY 3121 TGAACCTTCATGCTGTGCTGCTGCGCTTCAAAATGCGATCAATTAACGGATGAGTCAT 3180
Db 3121 TGAACCTTCATGCTGTGCTGCTGCGCTTCAAAATGCGATCAATTAACGGATGAGTCAT 3180
QY 3181 GGGGCCCCCTCTTGGGAGAGCTATTCTTGGCATTTGGGTGATGTTGCCATCTTATCCCT 3240
Db 3181 GGGGCCCCCTCTTGGGAGAGCTATTCTTGGCATTTGGGTGATGTTGCCATCTTATCCCT 3240
QY 3241 TTCTCAAGGTTGTTGGAAGGCAAAACAGAGACCAAGATTGTCATCGTCTGTCCA 3300
Db 3241 TTCTCAAGGTTGTTGGAAGGCAAAACAGAGACCAAGATTGTCATCGTCTGTCCA 3300
QY 3301 TTCTGCTGCTCAATCTTCTGCTCTCTTGGGTGCGATTGATCCTTCTTCCGGAAG 3360
Db 3301 TTCTGCTGCTCAATCTTCTGCTCTCTTGGGTGCGATTGATCCTTCTTCCGGAAG 3360
QY 3361 ATGATGTCGCTTCTTGAAGAGTGTGTTGATTGCACTAGAGTGCAGTGCATCAG 3420
Db 3361 ATGATGTCGCTTCTTGAAGAGTGTGTTGATTGCACTAGAGTGCAGTGCATCAG 3420
QY 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGCTGTTGTCCCATATTCAGTGT 3480
Db 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGCTGTTGTCCCATATTCAGTGT 3480
QY 3481 CTGTAGATAAGAGACATGAATGTCCCAAGTTCTTTGATCCATGCTGAACCTACTAA 3540
Db 3481 CTGTAGATAAGAGACATGAATGTCCCAAGTTCTTTGATCCATGCTGAACCTACTAA 3540
QY 3541 TATCTGAGAGATATCTGGGGGAAAATGAGGCTGCGGCAATCTTGTGACAGTTGGCCG 3600
Db 3541 TATCTGAGAGATATCTGGGGGAAAATGAGGCTGCGGCAATCTTGTGACAGTTGGCCG 3600
QY 3601 TGAATACAGCATATGCAAGTGTGATGTCAGACATCTTATTAATCTTGTGTCGACATA 3660
Db 3601 TGAATACAGCATATGCAAGTGTGATGTCAGACATCTTATTAATCTTGTGTCGACATA 3660
QY 3661 TAGATGGGCTGAGCCGAACAGCAAGTATTTTGTATCTGCACTGCTCCCGTGTACAACT 3720
Db 3661 TAGATGGGCTGAGCCGAACAGCAAGTATTTTGTATCTGCACTGCTCCCGTGTACAACT 3720
QY 3721 TGGTCTCAATAAGCAGGAGAGATGCATCTGCCAGTGAACAAGCAACCTGCACATT 3780

```

```

Db 3721 TGGTCTCAATAAGCAGGAGAGATGCATCTGCCAGTGAACAAGCAACCTGCACATT 3780
QY 3781 ATTTATGATGCTGTTCAATGAGGGCTGTTCAATTAAGTGTGCTATATACTAGAAA 3840
Db 3781 ATTTATGATGCTGTTCAATGAGGGCTGTTCAATTAAGTGTGCTATATACTAGAAA 3840
QY 3841 AACAGAAATTAACATTAATCTAAGTTAATTAAGTATGAATGCGCTGTTTGTGT 3900
Db 3841 AACAGAAATTAACATTAATCTAAGTTAATTAAGTATGAATGCGCTGTTTGTGT 3900
QY 3901 TGTGTAAGTATCATCTGAGTTGTTGTGAAAA 3936
Db 3901 TGTGTAAGTATCATCTGAGTTGTTGTGAAAA 3936

RESULT 4
AAZ58263
ID AAZ58263 standard; cDNA; 3776 BP.
XX
AC AAZ58263;
XX
DT 08-MAY-2000 (first entry)
XX
DE Corn cellulose synthase cDNA.
XX
KM Corn; maize; cellulose synthase; transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 2..3448
FT FT /+tag= a
XX FT /partial
XX
PN MO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15871.
XX
PR 14-JUL-1998; 98US-0092844.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
DR WPI; 2000-182431/16.
DR P-PSDB; AAY58832.
XX
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants -
XX
PS Claim 9; Page 37-38; 93pp; English.
XX
CC The present sequence is that of a contig of cDNA clones and encodes
CC a portion (see AAY58832) of corn cellulose synthase (CS). The cDNA
CC clones were isolated from corn cob, developing tassel, seedling,
CC shoot culture and leaf tissue cDNA libraries on the basis of
CC homology to Arabidopsis and cotton CS sequences. The invention
CC relates to isolated nucleic acid fragments encoding plant CS and to
CC CS polypeptides. It also relates to the construction of a chimeric
CC gene encoding all or a portion of the CS, in sense or antisense
CC orientation, where expression of the gene results in altered levels
CC of the CS in transformed host cells. The host cells can be used to
CC screen compounds for their ability to inhibit CS activity. CS
CC nucleic acids are also useful for producing transgenic plants having
CC altered levels of CS, and hence altered levels of fibre. CS may
CC also serve as a target for the development of novel herbicides.
XX
SQ Sequence 3776 BP; 931 A; 906 C; 999 G; 940 T; 0 other;

```

Query Match		65.9%;	Score 2592.6;	DB 21;	Length 3776;
Best Local Similarity		86.6%;	Pred. No. 0;		
Matches 2885;		Conservative	0;	Mismatches 439;	Indels 9; Gaps 2;
Qy	104	CTCTGCTTCGTGCGCGGAGTGAAGGGAGGAAGCGATGAGCGGAGCGCGGCT	163		
Db	139	CGCTTGCCCCGGGATCTGGAGATCTGGTAGCGCCAGGGGATGAGGCCAGCGCGGCT	198		
Qy	164	GGTGGCCGGCTCCCAACAACCGCAACGAGCTCGTGCATCCGCGCGAGCGGATCCCGG	223		
Db	199	GGTCGGCTCGCACAACCGGAACGAGCTCGTCATCCGCGGATGGCGAGCCAGG	258		
Qy	224	GCCGAAGCCGCGCGGAGCAGAAACGGGAGGTGTGCGCAGATTGGCGGCGAGCGAGTGG	283		
Db	259	GCCGAAGCCCATGGAACGAGCGGAACGGCAGGTGTGCCAGATTGGCGGCGAGCGTGGG	318		
Qy	284	CTTGCCCCCGCGCGGAGCCCTTCGTGGCGTGCAACGAGTGGCGCTTCCCGCTGCGCG	343		
Db	319	GCGCAACCCCGACGGGGAGCGCTCGTGGCCTGCAACGAGTGGCGCTTCCCATCTGCGCG	378		
Qy	344	GGACTGTCTACGAATACGAGCGCGGGAGGGCAACGAGAATGCCCCCACTGCAAGACTCG	403		
Db	379	GGACTGTCTACGAATACGAGCGCGCGGAGGGCAACGAGAATGCCCCCACTGCAAGACCCG	438		
Qy	404	ATACAAGCGCCTCAAGGGCTGCCAAGTGTGACCGGTGACGAGAGAGGACGGCGTGA	463		
Db	439	CTTCAAGCGCCTCAAGGGGTGCGCGCGCTGCGCGGAGCAGAGAGAGACGGCGTGA	498		
Qy	464	TGACCTGGCAACGAGTTCAACTGG--GACGGCCATGACTCGCAGTCTGTGCGCGAGTC	520		
Db	499	CGACCTGGAGAACGAGTTCAACTGGAGCACAACGAGTCCCACTCCAGTACTCGCGAGTC	558		
Qy	521	CATGCTCTACGGCCACATGAGCTACGGCGGTGAGGTGACCCCTAATGGCGCGCCACAAGC	580		
Db	559	CATGCTCCACGCCCAACATGAGCTACGGCGCGCGCGCGCCAGCTCGACGGCGTCCGCAAGC	618		
Qy	581	TTTCCAGCTCAACCCCAATGTTCCACTCTCAACCAACGGGCAATGCTGATGACATCCC	640		
Db	619	ATTCCACCCCATCCCAATGTTCCCTCTCTCAACCAACGAGACAGATGCTGACATCCC	678		
Qy	641	ACCGGAGCAGCAGCGCTGTGCTTTTCATGGGTGGGGGAAAGAGATACATCC	700		
Db	679	GCCGGAACCAAGACGCCCTTGTGCTCTGTCGTGGGTGGCGGGGAAAGAGATTCAACC	738		
Qy	701	CTTCCCTTATGCGGATCCAGCTTACCTGTGCAACCCAGGTCTATGGAACCCATCCAAGGA	760		
Db	739	TCTCCCGTACGGGATCCCAACCTTCCGTGTGCAACCGAGTCTATGGAACCTTCCAAGGA	798		
Qy	761	TCTTGCTGATATGGGTATGTAAGTGTCTTGGAAAGAACGATGGAGATTGGAAGCA	820		
Db	799	TCTGCGCCCATATGGCTACGGGAGCGTACATGGAAGAGAGATGGAAGCTGGAAGCA	858		
Qy	821	GAGACAAGAGAGATGCACCAAGCGGGAATGATGGTGGTGTATGATGTGACGATGC	880		
Db	859	GAAAGAGAGAGATGCACCAAGAGGAAAGATGGCGGCGGATGATGATGATGC	918		
Qy	881	TGATCTACCACTAATGATGAAGCAACAACCTGTCCAGAAAATTCACCTTCCATC	940		
Db	919	AGATCTACCACTAATGATGAAGCTAGACAGCCATTGTCCAGAAAAGATCCCGCTTCTTC	978		
Qy	941	AAGCCAGATTAACTCATATAGGATGATTATTCGCTTGTGGCTTGGGTTCTT	1000		
Db	979	AAGCCAAATCAACCCCTATAGATGATTATAATAATTCGGCTAGTGGTTTGTGTTCTT	1038		
Qy	1001	CTTCCACTACCGAGTGCATCCGGTGAATGATGCATTTGCTTGTGGCTCATATCTGT	1060		
Db	1039	CTTCCACTACCGAGTGCATCCGGTGGCTGATGCATTTGCTTATGGCTCATATCTGT	1098		
Qy	1061	TATCTGTGAATCTGTTTGGCATGTCTTGGATTCTTGATCAATTCCCAAGTGTTC	1120		
Db	1099	GATCTGTGAATTTGTTTGGCATGTCTTGGATTCTTGATCCAGTTTCCAAGTGTTC	1158		

Qy	1121	TATTGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTGCAACAAGAGGCCAGCC	1180		
Db	1159	TATCGAGAGGAAACCTATCTTGACCGGCTGAGTTAAGTTGACAAGAGAGGCATCC	1218		
Qy	1181	ATCTCACTTGCTCCAAATGATTCTTGTGACGTACGTTGATCCCTTAAGAACTCC	1240		
Db	1219	TTCTCACTCGCCCTGTGATTCTTGTGACGTACGTTGATCCCTTAAGAACTCC	1278		
Qy	1241	TTTGGTCAACAATACTGTTCTATCTATCTCTTCCGTTGATATCTCTGTTGATAAGT	1300		
Db	1279	ATTGGTCACTGCTAATACTGTTCTATCTATCTCTTCCGTTGATATCTCAGTTGATAAGT	1338		
Qy	1301	TTCTTGCTATGTTTCTGATGATGCTGCTCAATGCTTAAGTTGAAGCATTATCTGAAC	1360		
Db	1339	TTCACTGCTACGTTCTGATGATGCTGCTCCATGCTGAATTTGAAGCATTGTGAAAC	1398		
Qy	1361	ATCTGAATTTGCAAGAATGGGTTCTTCTGCAACGGTACAATATTGAACCTCGCGC	1420		
Db	1399	ATCTGAATTTGCAAGAATGGGTTCTTCTGCAAAAAGATATAGCCTTGAGCCTCGTGC	1458		
Qy	1421	TCCAGAGTGTACTTCCAAACAGAGATAGACTACTTGAAGAACAAGGTGCAGCAACTT	1480		
Db	1459	TCCAGAGTGTACTTCCAAACAGAGATAGACTACTTGAAGAACAAGGTGCAGCAACTT	1518		
Qy	1481	TGTTAGGAGAGAGAGCAATGAAGAGAGATAGAGATTCGAAGTGAAGATCAATGC	1540		
Db	1519	TGTTAGAGAACGAGAGCAATGAAGAGAGATAGAGATTCGAAGTCAAGATCAATGC	1578		
Qy	1541	CTTAGTTGCCAAAGCCCAAGAACTTCTGAAAGAGATGGAACAATGCAAGATGAACCC	1600		
Db	1579	CTTAGTTGCTAAAGCCCAAGAACTTCTGAGGAAGATGGAACAATGCAAGATGAACCTC	1638		
Qy	1601	CTGGCTGGAAACAATGTTCTGATCATCTCGAATGATTCAAGTCTTCTTGCCAAAG	1660		
Db	1639	ATGGCCGGAATATGTCCTGATCATCTCGAATGATTCAAGTTTCTTGCTGAAG	1698		
Qy	1661	CGGAGCCTTGACTGTGAGGAAATGAACCTGCCACGATTGGTTATGTTCTAGAGAA	1720		
Db	1699	TGGTGGCCATGATGTGAAGAAATGAGCTGCTCGATTGGTTATGTTCAAGAGAAA	1758		
Qy	1721	ACGACAGGCTATAACCATCATTAAGAAAGCTGTGCTATGAATGCATTGTCGAGTCTC	1780		
Db	1759	ACGACAGGCTATAACCATCATTAAGAAAGCTGTGCTATGAATGCATTGTCGAGTCTC	1818		
Qy	1781	TGCTGTACTAACCAATGCTTCCATATTGTTAACTTGGATTGTATCACTACATCAACA	1840		
Db	1819	TGCTGTACTAACCAATGCTTCTTATTGCTGAACCTGGATGTGATCACTATATCAATAA	1878		
Qy	1841	CAGCAAGGCTATAAAGGAACAATGTGTTTATGATGACCCTTACTAGAAAGAGGT	1900		
Db	1879	TAGTAAAGGCTATAAAGGAACAATGTGTTTATGATGATCCTTGTGAAAGAAAGT	1938		
Qy	1901	TTGCTATGTACAGTTCCCTCAAGATTGATGGGATGATCCCATGACCGATATGCTAA	1960		
Db	1939	TTGCTATGTACAGTTCCCTCAAGATTGATGGGATGATCCCATGATGATGCTAA	1998		
Qy	1961	CCGGAATGTTGCTTTTGTATATCAACATGAAAGTTTGATGTGATTCAGGGTCCAT	2020		
Db	1999	CAGAAATGTTGCTTTTGTATATCAACATGAAAGTTTGATGTGATTCAGGGGCCAAT	2058		
Qy	2021	TTATGTTGCTACTGATGTGTATTAGAAAGCAGGCAATTATGTTATGATGCCCCCA	2080		
Db	2059	TTATGTTGCTACTGATGTGTCTTCAAGAGCAGGCAATTATGTTATGCTAGATGCTCCA	2118		
Qy	2081	AACAAGAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGTGCTTGTGCTGTG	2140		
Db	2119	AACAAGAGCCACCATCAAGGACTTGCAACTGCTGGCCCAAGTGTGCTTGTGCTGTG	2178		
Qy	2141	CTGCTTTGGCAATAGAAAGCAAAAGAGACTACCAAAACCAAAACAGAGA-----AGAA	2194		
Db	2179	CTGCTTTGGTAAACAGAAAGCAAGAAAGAACCAAGACTCTAAACCTAAATTGAGAA	2238		
Qy	2195	AAAGTATTTATTTTCAAGAAAGAAAGAACCAATCCCTGCATATGCTCTTGTGAAAT	2254		

```

Db      | 2239 GATAAAGAAACTTTTAAAGAAAAAGAAAAATCAAGCCCCCTGCATATGCTCTTGTTGAAT 2298
Qy      | 2255 TGACGAAGCTGCTCCAGAGAGCTGAGAAATGAAAAAGCCGGTATTGTAAATCAACAAAATT 2314
Db      | 2299 TGATGAAGCCGCTCCAGAGAGCTGAAAAATGAAAAAGCTAGTATTGTAATCAACAGAAATT 2358
Qy      | 2315 AGAAAAAGAAATTTGGCCAACTCTTGTGTTTGTATCATCCACACTTCTCGAAGATGTGG 2374
Db      | 2359 GAAAAAGAAATTTGGCCAGTCTTCAAGTTTGTGTCATCCACACTTCTTGAGAAATGTGG 2418
Qy      | 2375 AACCTTGAAGAGTGAAGTCCCTGCTCTCTTTGAAAGAGCTATACATGTCATTAGTTG 2434
Db      | 2419 AACCTGAAGAGTGCAGTCCAGCTTCTCTTGAAGAGAGCTATACATGTCATTAGTTG 2478
Qy      | 2435 TGGTATGAAGACAGACAGACTGGGAAAAAGATGGCTGATCTATGATCAGTTAC 2494
Db      | 2479 TGGATATGAAGACAAACAGGCTGGGAAAAAGATATGGTTGATTTATGATCAGTCAC 2538
Qy      | 2495 AGAAGATATTTAACTGCTTTCAAGATGTCATGTTGGCGGTCAAATTTACTGCAT 2554
Db      | 2539 AGAAGATATTTCTTACTGGGTTTAAGATGACCTGCCATGGTTGGCGGTCAAATTTACTGCAT 2598
Qy      | 2555 ACCTAAACGGGTTGCATCAAAAGGTTCTGCACCTCTGAATCTTTCAAGATGCTTCACCA 2614
Db      | 2599 ACCTAAACGGGCCGCTTCAAAAGGTTCCGACCTCTCAATCTTCCGATGCTTCACCA 2658
Qy      | 2615 GGTGCTCGTGGGCTCTGGGCTTATGAGATCTTCTTCAAGCAATCATGCCCCCTTGG 2674
Db      | 2659 GGTCTTCGGTGGGCTCTGGTTCATGAAATTTCTTCAAGCAACCATGCCCCCTCTG 2718
Qy      | 2675 GTATGGGATGTGGCGGCTCGAAATTTTGAAGATTTTCTACATCACTCCATCGT 2734
Db      | 2719 GTATGGGATGTGGTGAAGTAAAGTTCCTGGAAGGTTTTCGTACATTAATCCATCGT 2778
Qy      | 2735 GTATCCTTGACATCTATTTCCCTCTTGCTTACTGTACATTTGCCCTGCTGTTATT 2794
Db      | 2779 ATACCTTGACATCTATCCCGCTCTTGCCCTATTTGACACATTTGCCCTGCTGCT 2838
Qy      | 2795 GACAGGAAATTTATCACTCCAGAGCTGAATATGTGCCAGCCCTGTGCTCATGTCACT 2854
Db      | 2839 GACAGGAAATTTATCAAGCCAGAGCTTAACAAATGTGCCAGCCCTGTGCTCATGTCACT 2898
Qy      | 2855 TTTATCTGCATTTTGTCTACAGCATCTCTAGAAATGAGATGAGTGTGTTGAATGA 2914
Db      | 2899 TTTCATCTGCATTTTGTCTACAGCATCTCTGAATGAGATGAGTGTGTTGAATGA 2958
Qy      | 2915 TGACTGTTGAGGAATGACAGATTTCTGGGTCATTTGAGAGTGTCTCTCACACTCTTTGC 2974
Db      | 2959 TGACTGTTGAGGAATGACAGATTTTGGGTCATTTGAGAGGCGGTCTTTCACATCTCTTGC 3018
Qy      | 2975 TGTGTTCCAGGGACTTCTCAAGGTCATAGCTGGTGTGATACAAAGCTTCAACCGTGACATC 3034
Db      | 3019 TGTGTTCCAGGGACTCTCAAGGTCATAGCTGGTGTGACACAGAGCTTCACTGTGACATC 3078
Qy      | 3035 AAAGGGTGAGATGATGAGAGGTTCTCAGAGCTATATACATTCAAATGGACTACCTTATT 3094
Db      | 3079 CAAGGCGGAGAGAGAGAGGTTCTCAGAGCTGTACACATTCAATGAGAGCAACCTTCT 3138
Qy      | 3095 GATACCTCTACCACTTGCTTCTATGAACCTTCAATGCTGTGTGCGTGGCGTTTCAAA 3154
Db      | 3139 GATACCTCCGACCACTCTCTCTACTGAACCTTCAATGAGTGTAGTGGATCTCCAA 3198
Qy      | 3155 TGGCATCAATAACGATATGAGTATGAGTGGGCCCCCTTTGGGAAGCTATCTTGCATT 3214
Db      | 3199 TGGCATCAACAACGATATGATGAGTGGGCCCCCTGTTCGGGAAGCTTCTTTCATT 3258
Qy      | 3215 TTGGGTGATGTCATCTTTATCCCTTCTTCAAAAGGTTGGTTGGAAGGCAACAGAGAC 3274
Db      | 3259 TTGGGTGATGTCATCTTTATCCCGTCTCTCAAGGCTGTGTTGGAGGCAACAGAGAC 3318
Qy      | 3275 ACCAAGCATGTCATGCTGTGCTCATTTGCTGCTTCAATCTTCTGCTCTTGGGT 3334

```

```

Db      | 3319 GCCAAGATGTCATTTGCTGTGTCATCTCTGCTTCCATCTTCTGCTTGCATCTTCTGCTTGGGT 3378
Qy      | 3335 TCGATGATCCTTTCTCTTCCGAGAGATGATGTCCTCGCTTCTTGAGAGAGTGTGTTGA 3394
Db      | 3379 CCGATGACCCGTTCTCTTCCGAGAGATGATGTCCTCGTGTGAGAGAGTGTGTTGA 3438
Qy      | 3395 TTGCACTAGATGTCAAGTGCATCAGCTCCCC 3427
Db      | 3439 TTGCACTAGAGAGTCAAGCAGTGAATTTCCC 3471

RESULT 5
AAZ99491
ID      AAZ99491 standard; DNA; 3568 BP.
XX
AC      AAZ99491;
XX
DT      03-JUL-2000 (first entry)
XX
DE      DNA encoding a maize cellulose synthase.
XX
KW      Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW      transgenic plant; plant breeding marker; ss.
XX
OS      Zea mays.
XX
FH      Key
FT      CDS
FT      63..3236
FT      /tag= a
FT      /product= "cellulose synthase"
FT      /note= "no termination codon given"

XX      WO200009706-A2.
PN      24-FEB-2000.
XX
XX      16-AUG-1999; 99WO-US18760.
PF      17-AUG-1998; 98US-0096822.
PR
XX      (PION-) PIONEER HI-BRED INT INC.
PA
XX      Dhuga KS, Helentjaris TG, Bowen BA, Wang X;
PI
XX      WPI; 2000-224343/19.
DR      P-PSDB; AAY84107.
DR
XX
PT      New genes which encode maize cellulose synthase polypeptides in plants
PT      useful for modulating the expression of cellulose synthase in plants
PT      and to produce transgenic plants expressing the novel protein -
XX
XX      Claim 1; Page 82-87; 119pp; English.
PS
XX
CC      The present sequence encodes a maize cellulose synthase polypeptide.
CC      The cellulose synthase can be used for the improvement of stalk quality
CC      for improved stand or silage. It also provides an increased concentration
CC      of cellulose in the pericarp, hardening the kernel and improving its
CC      handling ability. The sequences are used to produce transgenic plants
CC      and seeds expressing the cellulose synthase. The polynucleotide is
CC      used for modulating, preferably increasing, the level of the synthase
CC      in a plant cell. The plants are preferably monocots. The polynucleotide
CC      is also used as a probe or primer in the detection quantitation or
CC      isolation of gene transcripts. The probes are useful in detecting
CC      deficiencies in the level of mRNA in screenings for desired transgenic
CC      plant, for detecting mutations in the gene, for monitoring upregulation
CC      of expression or changes in enzyme activity in screening assays of
CC      compounds, for detection of any number of allelic variants of the gene,
CC      or for use as molecular markers in plant breeding programs. The
CC      isolated nucleic acids of the present invention can also be used for
CC      recombinant expression of their encoded polypeptides or for use as
CC      immunogens in the preparation and/or screening of antibodies. The
CC      proteins can be employed in assays for enzyme agonists or antagonists
CC      of enzyme function or for use of immunogens or antigens to obtain

```



CC antibodies specifically immunoreactive with a protein.

Sequence 3568 BP; 879 A; 825 C; 941 G; 922 T; 1 other;

Query Match	64.74;	Score 2546.8;	DB 21;	Length 3568;
-------------	--------	---------------	--------	--------------

```

Best Local Similarity 86.86; Pred. NO. 0;
Matches 2829; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

```

Matches 2829; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

QY	177	CACAAACCGCAACGAGCTCGTGCATATCCGCCGACGCGCATCCCGGGCCGAAGCCGCCG	236
Db	6	CCCAACGCGTCCGAGCTCGTGCATATCCGCCGATGCGAGCCAGGCGCGAAGCCCATG	65
QY	237	CGGAGCAGAAACGGGAGGTGTGCAGATTGCGCGACGACGTCGCGCTTGCCCCCGGC	296
Db	66	GACCAGCGGAACGCGCAGGTGTGCAGATTGCGCGACGACGATGGGGCGCAACCCCGAC	125
QY	297	GGGACCCCTTCTGTGCGTGCACGAGTGCCTTCCCGCTGCGCGGACTGCTACGAA	356
Db	126	GGGAGCCCTTTCGTGCGCTGCACGAGTGCCTTCCCATCTGCGGGACTGCTACGAG	185
QY	357	TACGAGCGCCGGGAGGSCACGCAACTGCCCCAGTGCAGACTCGATTACAAGCGCTC	416
Db	186	TACGAGCGCCCGGAGGSCACGCAACTGCCCCAGTGCAGACTCGCTTCAAGCGCTTC	245
QY	417	AAGGGCTGCCAACGTGTGACCGGTGACGAGAGGAGGACGGCGTCGATGACCTGGACAAC	476
Db	246	AAGGGGTGCGCGCGTGCCTCGGGGACGAGAGGAGACGGCGTCGACGACCTGAGAAC	305
QY	477	GAGTCAACTGG---GACGGCCAATGACTCGAGTCTGTGCGCGAGTCCATGCTACGGC	533
Db	306	GAGTCAACTGGAGCGACAGACGACTCCAGTACTCGCGAGTCCATGCTCCACGCC	365
QY	534	CACATGAGCTACGCGCGTGAAGTGACCTAATGCGCGCCACAAGCTTTCAGCTCAAC	593
Db	366	CACATGAGCTACGCGCGCGCGCCGACCTCGACGGCGTCCGCGACCATTCCACCCCATC	425
QY	594	CCCAATGTTCCACTCTCTCACCAACGGGCAATGTTGATGACATCCCAACGGAGCAGCAC	653
Db	426	CCCAATGTTCCCTCTCTCACCAACGGACAGATGTCATGACATCCCGCGGACCAGCAC	485
QY	654	GCGGTGTCCTTCTTTCATGGGTGTGGGGGAAAGAGATACATCCCTTCTTATGCG	713
Db	486	GCCCTGTGCCCCCTCTGTGTGGGTGGCGGGGAAAGAGATTCACTCTCCGTACGCG	545
QY	714	GATCCAGCTTACCTGTGCAACCCAGTCTATGGAACCCATCCAGATCTTGCTGCATAT	773
Db	546	GATCCCAACCTTCTCTGTGCAACCGAGTCTATGGAACCTTCCAGATCTCGCCGATAT	605
QY	774	GCGTATGTTAGTGTGCTTGGAAAGAACGGATGGAATTTGAAACAGACAGAGAGAG	833
Db	606	GCGTACGGGAGCGTAGCATGGAAGAGAGGATGGAGAGCTGGAAGCAGAGAGAGAG	665
QY	834	ATGCAACCAAGACGGGGAATGATGTTGTGTGTGATGATGTTGATGCTGATCTACCACTA	893
Db	666	ATGCAACCAAGACGAGGAACGATGGCGGCGGATGATGTTGATGATGATCTACCACTA	725
QY	894	ATGATGAAGCAAGACAACAACCTGTCCAGGAAATTCCACTTCATCAAGCCAGATTAAAT	953
Db	726	ATGATGAAGCTAGACAGCCATTGTCCAGAAAGATCCCGCTTCTTCAAGCCAAATCAAC	785
QY	954	CCATATAGGATGATTATCATTAATCGGCTTGTGTTTGGGGTCTTCTTCCACTACCGA	1013
Db	786	CCCTATAGGATGATTATTAATAATTGGCTAGTGTGTTTGTGTTCTTCTTCCACTACCGA	845
QY	1014	GTCATGATCCCGGTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGTGAATC	1073
Db	846	GTCATGATCCCGGTGCTGATGATTTGCTTATGCGCTCATATCTGTGATCTGTGAATT	905
QY	1074	TGCTTTGCCATGTCTTGGATTCTTGATCAATTCCCAAGTGTTCCTATTGAGAGAGAG	1133
Db	906	TGCTTTGCCATGTCTTGGATTCTTGACCAAGTTTCCAAAGTGTTCCTATCGAGAGGGA	965
QY	1134	ACTTACCTAGACCGGCTGTCACTGAGGTTTCACAAGAGAGCCAGCATCTCAACTTGCT	1193

[illegible]

Db 2046 TTTAAGAAAAAGAAATCAAGCCCCGTCATATGCTCTTGGTGAATGTATGAAGCCGCT 2105  
QY 2268 CCAGAGCTGGAATGAAAAAGCCGGTATTGTAATCAACAAAAATTAGAAAAAATT 2327  
Db 2106 CCAGAGCTGAAAAATGAAAAAGCTAGTATTGTAATCAACAGAGTTGAAAAAATT 2165  
QY 2328 GGCCAATCTTCTGTTTTTTGTACATCCACACTTCTCGAAGATGGTGAACCTTGAAGAGT 2387  
Db 2166 GGCCAGTCTTCAGTTTTTTGTGATCCACACTTCTTGAGAAATGGTGAACCCCTGAAGAGT 2225  
QY 2388 GCAAGTCTGCTTCTCTTTTGAAGAAGCTATACATGTCAATTAGTTGTTATGAAGAC 2447  
Db 2226 GCCAGTCCAGTCTCTCTTGGAAGAAGCTATACATGTCAATTAGTTGATATGAAGAC 2285  
QY 2448 AAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCAGTTACAGAAAGATATTCTA 2507  
Db 2286 AAAACAGGCTGGGAAAAAGATATTGGTTGATTTATGATCAGTCACAGAAAGATATTCTT 2345  
QY 2508 ACTGTTTCAAGATGCATTGTCAATGGTGGCGGTCAATTACTGCATACCTAAACGGGTT 2567  
Db 2346 ACTGGGTTTAAAGTCACTGCCATGGTGGCGGTCAATTACTGCATACCTAAACGGGCC 2405  
QY 2568 GCATTCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTACCGAGTCTTCGGTGG 2627  
Db 2406 GCCTTCAAAGGTTCCGACCTCTCAATCTTCCGATCGTTTACCCAGGTTCTTCGGTGG 2465  
QY 2628 GCTCTTGGGCTATGAGATCTTCTTCAGCAATCATTGCCCTCTTGGTATGGGTATGGT 2687  
Db 2466 GCTCTTGGTCAATGAAATTTTGTTCAGCAACCACTGCCCTCTTGGTATGGGTATGGT 2525  
QY 2688 GCGGCTGGAATTTTGGAAAGATTTTCCATCATCACTCCATCGTATCCCTTGAGACA 2747  
Db 2526 GGTGACTAAAGTTCCTGGAAAGTTTTCGATACATTACTCCATCGTATACCTTGAGACA 2585  
QY 2748 TCTATTCCTCTTGGCTTACTGTACATTCGCTGCCATCTGTTATTCACAGGAATTT 2807  
Db 2586 TCTATCCCGCTCTTGGCCTATTGCACATTCGCTGCCATCTGCTTGCCTGACAGGAATTT 2645  
QY 2808 ATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGTTTCATGTCTACTTTTATCTGCATT 2867  
Db 2646 ATCAGCCAGAGCTTAACAATGTTGCCAGCCTGTGTTTCATGTCTACTTTTTCATCTGCATT 2705  
QY 2868 TTGCTACGAGCATCTAGAAATGAGATGAGTGTGTTGAATTGATGACTGTGGAGG 2927  
Db 2706 TTGCTACGAGCATCTGAAATGAGATGAGTGTGTTGAATTGATGACTGTGGAGG 2765  
QY 2928 AATGAGCAGTTCGGGTCAATTGGAGGTGTCTTCACACCTCTTGTGTTTCCAGGGA 2987  
Db 2766 AACGAGCAGTTCGGGTCAATTGGAGGTGTCTTCACATCTCTTGTGTTTCCAGGGA 2825  
QY 2988 CTTCTCAAGTCATAGCTGTGTTGATACAACTTACCCTGACATCAAGGGTGAGAT 3047  
Db 2826 CTCCTCAAGTCATAGCTGTGTTGATACAACTTACCCTGACATCAAGGGTGAGAT 2885  
QY 3048 GATGAGAGTTCAGAGCTATATACATTCAAATGAGCTACCTTATGATACCTCTTACC 3107  
Db 2886 GACGAGGAGTTCAGAGCTGTACACATTCAAATGAGCTACCTTATGATACCTCTTACC 2945  
QY 3108 ACCCTGCTTCTATGAACTTCATTTGAGTGTGCTGGCGTTCAATGCGATCAATAAC 3167  
Db 2946 ACCCTGCTTCTATGAACTTCATTTGAGTGTGCTGGCGATTCATGCGATCAATAAC 3005  
QY 3168 GGATATGATGATGAGGCCCCCTCTTGGAGAGCTATTCTTTCATTTGGGTGATGTC 3227  
Db 3006 GGATATGATGATGAGGCCCCCTCTTGGAGAGCTATTCTTTCATTTGGGTGATGTC 3065  
QY 3228 CATCTTATCCCTTCTCAAGGTTTGGTGAAGGCAAAACAGAGCAACGATGTC 3287  
Db 3066 CATCTTATCCCTTCTCAAGGTTTGGTGAAGGCAAAACAGAGCAACGATGTC 3125  
QY 3288 ATGCTGTGTCATTTCTGCTGCTTCAATCTTCTGCTCTTGGGTGCGATGATCCT 3347  
Db 3126 ATGCTGTGTCATTTCTGCTGCTTCAATCTTCTGCTCTTGGGTGCGATGATCCT 3185

QY 3348 TTCCTGCGAAGATGATGTCCTCTTTCGAGAGAGTGTGTTGATTCGAAGTATG 3407  
Db 3186 TTCCTGCGAAGATGATGTCCTCTTTCGAGAGAGTGTGTTGATTCGAAGTATG 3245  
QY 3408 GTCAGTGCATCAGCTCCCC 3427  
Db 3246 GTCAGCAGCTGACTTCCCC 3265

RESULT 6  
AAZ58265  
ID AAZ58265 standard; cDNA; 3786 BP.  
XX  
XX AC AAZ58265;  
XX  
DT 08-MAY-2000 (first entry)  
XX  
DE Corn cellulose synthase cDNA clone p0097.cqrad17rc.  
XX  
KM Corn; maize; cellulose synthase; transgenic plant; ss.  
XX  
OS Zea mays.  
FH Key location/Qualifiers  
FT CDS 2..3499  
FT /\*tag= a  
FT /partial  
XX  
PN WO200004166-A2.  
XX  
PD 27-JAN-2000.  
XX  
PF 13-JUL-1999; 99WO-US15871.  
XX  
PR 14-JUL-1999; 98US-0092844.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JB, Miao G;  
PI Rafalski JA, Thorpe CJ;  
XX  
DR WPI; 2000-182431/16.  
DR P-PSDB; AAY58834.  
XX  
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
PT probes for isolating cDNAs and genes encoding homologous proteins, for  
PT producing transgenic plants -  
XX  
PS Claim 9; Page 44-45; 93pp; English.  
XX  
CC The present sequence is that of cDNA clone p0097.cqrad17rc encoding  
CC a portion (see AAY58833) of corn cellulose synthase (CS). The cDNA  
CC clone was isolated from a european corn borer 4 times-infected corn  
CC stage V9 whorl section cDNA library on the basis of homology to  
CC Arabidopsis and cotton CS sequences. The invention relates to  
CC isolated nucleic acid fragments encoding plant CS and to CS  
CC polypeptides. It also relates to the construction of a chimeric  
CC gene encoding all or a portion of the CS, in sense or antisense  
CC orientation, where expression of the gene results in altered levels  
CC of the CS in transformed host cells. The host cells can be used to  
CC screen compounds for their ability to inhibit CS activity. CS  
CC nucleic acids are also useful for producing transgenic plants having  
CC altered levels of CS, and hence altered levels of fibre. CS may  
CC also serve as a target for the development of novel herbicides.  
XX  
SQ Sequence 3786 BP; 846 A; 967 C; 1111 G; 862 T; 0 other;

Query Match 56.9%; Score 2241; DB 21; length 3786;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;  
QY 123 GGAAGTGAAGGGGAGAGGAGATGAGGCGAGCGCGGCTGTGGCGGCTCCACAAC 182

Db 194 GGGCGGGGGGGGGCTGAGATGAGGCTAGCGGGGCTGTGGCCGCTCGCATAC 253  
Oy 183 CGCAACGAGCTCGTCGTCATCCGCGCGAGCGCGA-TCCCGGGCCGAAGCCCGCGGGA 241  
Db 254 CGGAACGAGCTGTGTGATCCGCGCGGAGCTGGGAGCCCGCGGGCGCGCGG 313  
Oy 242 GCAGAACGGGC-----AGGTGTCCAGATTGGCGGCGAGCAGCTGGCCCTTGCCCCGGC 296  
Db 314 GCGCGCCGGCGGAGGCGCGCTGCGAGATATGCGGCGAGAGGTGGGGTGGCTTGAC 373  
Oy 297 GGGGACCCCTTCGTGGCGGTGCAACGAGTGGCCTTCCCGCTGCGCGGAGCTGTACGAA 356  
Db 374 GGGGAGCCCTTCGTGGCGGTGCAACGAGTGGCCTTCCCGCTGCGCGCTGTACGAG 433  
Oy 357 TACGAGCCCGGGAGGGCGACGAGAACTGCCCCAGTGCAAGACTCGATACAAGCGCTC 416  
Db 434 TACGAGCGCGCGAGGGCTCGCAAGCGTGGCCGAGTGCAAGACCGGCTACAAGCGCTC 493  
Oy 417 AAGGCTGCCACGTGTGACCGGTGACGAGAGGAGGACGGCGCTGATGACCTGACAAAC 476  
Db 494 AAGGCTGCCCGCGGGGTGGCCGCGAGGAGGAGGACGGCGTGAAGACCTGAGGGGC 553  
Oy 477 GAGTTCAACTGG-----GACGGCCATGACTCGCAGTCTGTGGCCGAGTCC 521  
Db 554 GAGTTCGGCTGACGAGGACGGCGCGCGCCGACGAGGACGACCCGCAATACGTGCGGAGTCC 613  
Oy 522 ATGCTCTAGGCGCACATGAGCTAGCGCGGTGAGGTGACCCCTAATGCGCGCCACAAGCT 581  
Db 614 ATGCTCAGGCGCGAGATGAGCTACGGCGCGCGCGCGA-----CGCGCACCCCGGC 664  
Oy 582 TTCCAGCTCAACCCCAATGTTCCACTCCTCACCAACGGGCAATGTGGATGACATCCCA 641  
Db 665 TTCAAGCCCGTCCCCAACGTGCGGCTCTCACCAACGGCCAGATGTTGATGACATCCCG 724  
Oy 642 CCGAGACGACGCGCTGTGCTCTTTCATG-----GTTGGTGGGGAAGAG 692  
Db 725 CCGGAGCAGCAGCGGCTGTGCTCTTACATGAGCGCGCGCGCGGGGGAAGAG 784  
Oy 693 ATACATCCCTTCCTTATGCGGATCCCACTTACCTGTGCAACCCAGGTCTATGACCCA 752  
Db 785 ATCCACCGCTCCCTTTCGAGATCCCACTTCAGTGCAACCGAGATCCATGAGCCCG 844  
Oy 753 TCCAAGATCTGTGTCATATGGTATGTGTGCTTGAAGAAACGATGAGAAAT 812  
Db 845 TCCAAGATCTGGCCGCTACGATATGCGAGCGTGGCTGGAAGAGAGATGAGGCG 904  
Oy 813 TGAAGCAGAGACAAGAGAGATGACACAGCGGGAATGATGCTGTGTGAT--GAT 869  
Db 905 TGAAGCAGAGAGAGAGCGGCTGCGAGATGTCAGAGCGAGGGTGGCGGTGATGGGAT 964  
Oy 870 GGTGACGATGCTGATCTACCACTAATGATGAAGCAAGACAACACTGTCCAGAAAATT 929  
Db 965 GGGCAGCATGAGATCTGCCACTAATGATGAAGCTAGGCAGCACTGTCCAGAAAAGTTC 1024  
Oy 930 CCACCTTCATCAAGCCAGATTAATCATATAGATGATTATCATTTGCGCTGTGTT 989  
Db 1025 CTAATATCATCAAGCCGAATTAATCCTACAGATGATTATCGTATCCGGTGTGTGTT 1084  
Oy 990 TTGGGTTCTTCTTCCACTACCGAGTATGATCCGGTGAATGATGATTTGCTTGTG 1049  
Db 1085 TTGGGTTCTTCTTCCACTACCGAGTATGATCCGGGAAGATGATTTGATTTGTG 1144  
Oy 1050 CTCAATCTGTATCTGTGAATCTGTGGCATGTCTTGATTTCTTGATCAATTCCCA 1109  
Db 1145 CTCAATCTGTATCTGTGAATCTGTGGTGTGGCATGTCTTGATCAATTCCCA 1204  
Oy 1110 AAGTGTCCCTATTGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGTTGACAAAG 1169  
Db 1205 AAGTGTCTTCAATCGAGAGAGACTTAACCTGAGCCGTTGTCACTAAGGTTGACAAAG 1264  
Oy 1170 GAAGGCGAGCCATCTCAACTTGTCTCAATTGATTTCTTGTGAGTACGGTTGATCCCTTA 1229

Db 1265 GAAGTCAACCCCTCTCAGCTTGTCCAAATGACTTCTTTGTCAGTACGGTTGATCCACA 1324  
Oy 1230 AAGAACCTCTCTTGTGTCACAACAATACTGTTCTATCTAATCTTTGCGGTGATTAATCT 1289  
Db 1325 AAGAACCTCTCTTGTGTCAGAGCGAAACACTGTCTTCCATCTTTCTGTGATTAATCCG 1384  
Oy 1290 GTTGATTAAGTTTCTGTATGTTTCTGATGATGTGTGCAATGCTAACCTTGAAGCA 1349  
Db 1385 GTTGAGAAAGTCTCTCTGTATGTTTCTGATGATGTGTGCAATGCTTAAGTTGAAGCA 1444  
Oy 1350 TTATCTGAAAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAAAACGTTACAAATTT 1409  
Db 1445 TTGTCTGAAAACATCTGAATTTGCAAGAAATGGGTTCTTTCAAGCAAAAGTTTAATATC 1504  
Oy 1410 GAACCTCGCGCTCCAGAGTGTACTTCCAAACAGAAATAGACTACTTGAAGACAAAGTG 1469  
Db 1505 GAGCTCTGTCTCTGAGTGTACTTCCAAACAGAAATAGACTACTTGAAGACAAAGTG 1564  
Oy 1470 GCAGCAAACTTGTAGGAGAGAGAGCAATGAAGAGAGAGTATGAGAAATTCAGGTG 1529  
Db 1565 GCTGCTTCAATTTGTAGGAGAGAGAGGCGATGAAGAGAGATACGAGAAATTCAGGTG 1624  
Oy 1530 AGAATCAATGCTTACTTGTCCAAAGCCCAAGAAAGTTCTGAAAGAGATGACAAATGCA 1589  
Db 1625 AGAATCAATGCTTGTGTCAAAGCCCAAGAAAGTTCTGAGGAAAGATGACAAATGCA 1684  
Oy 1590 GATGAAACCCCTGCGCTGGAACCAATGTTCTGTATCATCTTGAATGATTCAGGTCTTC 1649  
Db 1685 GATGAAAGCCCTGCGCTGGAACCAACGTACGCGATCATCTGGAATGATTCAGGTATTC 1744  
Oy 1650 CTTGGCCAAAGCGAGCGCTTGACTGTGAGGGAATGAATGCCACGATTTGTTATGTT 1709  
Db 1745 CTTGGCCAAAGTGGCGGTGTGATGTGAAGGAATGATTTGCCCTGCTGTTATGTC 1804  
Oy 1710 TCTAGAGAGAAACGACGAGCTAATAACCATCATTAAGAAAGCTGTGCTATGAATGCTG 1769  
Db 1805 TCGAGAGAAAGAGCGCAGTTAATAACCATCAAGAAAGCTGTGCCATGAATGCACTG 1864  
Oy 1770 GTCCGAGTCTGTGCTACTAACAATGCTCCATATTTGTTAACTTGATTTGATGATCAC 1829  
Db 1865 GTCCGAGTCTGTGCTACTAACAATGCTCCATATTTGAATTTGATGATGATGATCAC 1924  
Oy 1830 TACATCAACAACAGCAAGGCTAATAAGGAAGCAATGTGTTTATGATGACCCCTTAATA 1889  
Db 1925 TACATCAACAATAGCAAGGCTAATAAGAGGCTATGTGTTTCAATGATGATCTTGTG 1984  
Oy 1890 GGAAGAAAGTTGTCTATGTACAGTTCCTCAAAAGTTGATGGGATGATGCCATGAC 1949  
Db 1985 GGAAGAAAGTGTGTATGTACAGTTCCTCAGAGTTGATGTGATGACAAAATGAT 2044  
Oy 1950 CGATATGCTAACCGGAATGTTGCTTTTGTGATCAACATGAAGGTTTGATGTATTT 2009  
Db 2045 CGATATGCTAACAGGAACGTGCTTTTGTGACATCAATGAAGGTTTGACGCTATTT 2104  
Oy 2010 CAGGGTCCAAATTTATGTGTACTGATGTGTATTTAGAAAGGCAAGCATTAATATGTTAT 2069  
Db 2105 CAAGAACCCATTTATGTGTGTACTGATGTGTTCAGACGGCAAGCATGTATGTAT 2164  
Oy 2070 GATGCCCCCAAAACAAGAACCCACCATCAAGGACTTGCAACTGTGGCCCAAGTGTGC 2129  
Db 2165 GATGCTCTTAAACGAAGAACCCACCATCAAGAACTTGCAACTGTGGCCCAAGTGTGC 2224  
Oy 2130 TTTTGTGTTGTGCTTGTGGCAATAGGAAGCAAAAGAACTAACAAACCAAAACAGAG 2189  
Db 2225 CTCTCTGTGTGTCAGAGCAAGAACAAATTAAGAAAGAACTAACAAACCAAGCGAG 2284  
Oy 2190 AAGAAAAAGTTATTTTCAAGAAAGAAAGAAACCAATCCCGCATATGCTTGTG 2249  
Db 2285 AAGAAAGAAAGATTTTTCAGAAAGCAGAAACCCATCTCCGCAATATGCTTGGGT 2344  
Oy 2250 GAAATTGACGAAGCTGCTCAGAGAGCTGAAGATGAAGAAAGCCGATTTGTAAATCAACA 2309  
Db 2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAGAAAGCCGAATCGTAATCAACAG 2404



QY	2310	AAATTGAAAAAATTGGCCAAATCTTCTGTTTGTATCATCCACACTTCTCGAAGAT	2369
Db	2405	AAACTAGAGAAGAAATTGGGCAGTCTTCTGTTTGTTCGATCAACACTTCTTGGAAC	2464
QY	2370	GGTGAACCTTGAAGATGCAAGTCTGCTTCTCTTTGAAAGAAGCTATACATGTCATT	2429
Db	2465	GGAGGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTTGAGGAAGCTATACATGTTATC	2524
QY	2430	AGTTGTGTTATGAAGACACACAGACTGGGAAAAAGATTTGGCTGATCTATGATCA	2489
Db	2525	AGCTGCGGCTACGAAGACACAGACCGACTGGGAAAAAGATTGGCTGATTTACGGATCG	2584
QY	2490	GTTACAGAGATATTCTAACTGTTTCAAGATGATTCATGTTGGCGGTCAATTAC	2549
Db	2585	ATCACAGAGATATCTGACTGGAATTTAAGATGACATGCTGCGGTCTAATTAC	2644
QY	2550	TGCATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTCGAAATCTTTCAGATCGTCT	2609
Db	2645	TGCATCCCCGAAGCGGCTGCAATTCAAAGGTTCTGCGCTCGAACCTTTCGACCGTCTT	2704
QY	2610	CACCAGGTGCTTCGGTGGGCTTTGGGCTATTGAGATCTCTTCAGCAATCATGCGCT	2669
Db	2705	CACCAGGTCTTCGGTGGGCTTTGGGCTCGTCCGTCGAATTTCTTCAGCAAGCACTGCCA	2764
QY	2670	CTTTGGTATGGGTATGTTGGCGGCTCGAAATTTTGGAAAGATTTTCTCATCATCTCC	2729
Db	2765	CTTTGGTACGGATACGGCGGGGCTAAATTCCTGGAAGGTTTCTTATATCAACTCC	2824
QY	2730	ATCGTATCCTTGACATCTATTCCTCTTGCTTACTGTACATTTGCCCTGCATCTGT	2789
Db	2825	ATCGTTTATCCCTGAGCGTCAATCTCTCTGCTTACTGTACTTGCCTGCCATCTGC	2884
QY	2790	TTATTGACAGGAAATTTATCACTCCAGAGCTGAATAAATGTTGCCAGCCTGTGGTTCATG	2849
Db	2885	CTGCTCACGGGGAAGTTTATCACACAGAGCTTACCAATGTCGCCAGTATCTGGTTCATG	2944
QY	2850	TCACTTTTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGGAGTGTGTTGA	2909
Db	2945	GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAATGAGGTGAGTGGCGTGGCC	3004
QY	2910	ATTGATGACTGGTGGAGGATGAGCAGTTCCTGGGTCAATTGAGAGGTGTCTCTCACACTC	2969
Db	3005	ATCGACGACTGGTGGAGGACGAGCAGTTCCTGGGTCACTGAGAGGCGTTTCGGCGCATCTG	3064
QY	2970	TTTGCTGTGTTCCAGGACTTCTCAAGTCATAGCTGTGTGATACAAGCTTCAACCGTG	3029
Db	3065	TTCCGGGTGTTCCAGGGCGCTGTGAAGTGTTCGCGGCATCGACAGACTTCACCGTG	3124
QY	3030	ACATCAAGGGTGAGATGATGAGAGGTTCTCAGAGCTATATACATTCAATGAGACTCAC	3089
Db	3125	ACGTGGAAGCGCGGGGACGACGAGAGTTCCTCGGAGCTGTACACGTTCAAGTGACCAAC	3184
QY	3090	TTATTGATACCTCCTACCACTTGTCTTATTTGAACCTTCAATTTGGTGTGCTGGCGTT	3149
Db	3185	CTGCTGATACCCCCGACACGCTCTCTCTGTAACCTTCAATGGGGTGTGGCCGGGATC	3244
QY	3150	TCAAAATGCGATCAATAACGATATAGTCATGGGGCCCCCTCTTTGGGAAGCTATTTCTT	3209
Db	3245	TCGAACGCGATCAACAACGGGTACAGTCGTGGGGCCCCCTGTTCCGGAAGCTCTTCTTC	3304
QY	3210	GCAATTTGGGTGATTTGCCATCTTATCCCTTCTCAAGGTTGTTGGAAGGCAAAAC	3269
Db	3305	GCCTTCTGGGTGATCTGCACCTGTACCCGTTCTCAAGGGTCTGGTGGGAGGCAAGAC	3364
QY	3270	AGGACACCAACGATTTGTCATCTGTCTGTCATTTCTGTGGCTTCAATCTTCTCGCTCTT	3329
Db	3365	AGGACGCCGACGATCTCATCTGTCTGTCATCTCTGTGGCTCGAATCTTCTCGCTCTTG	3424
QY	3330	TGGGTTCCGATTGATCTTTTCTTTGCGAAGGATGATGTCCTTCTTGAGGAGTGTGGT	3389
Db	3425	TGGGTCCCGCTGCACCCGTTTCTTCGCAAGACCAACGCGCCCTCTCGAGGAGTGTGGC	3484

Query Match	Best Local Similarity	Score	DB	Length
56.9%;	81.2%;	2241;	21;	3813;
		Pred No. 0;		

  

Query Match	Best Local Similarity	Score	DB	Length
56.9%;	81.2%;	2241;	21;	3813;
		Pred No. 0;		

Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;					
Qy	123	GGAGTGAAGGGAGGAAGCGATGAGGCGAGCGCGGCTGTGTGCGCGCTCCACAAC	182		
Db	194	GGGCGGGGGGGGGGGCTGTAGATGAGGCTAGCGGGGCTGTGTGCGCGCTCGCATAAC	253		
Qy	183	CGCAACGAGCTCGTCATCCGCGCGCACGCGCA-TCCCCGGCGGAAGCCCGCGGGA	241		
Db	254	CGGAACGAGCTGTGTGATCCGCGCGCAACCGGAGTCGGGAGCCGCGGCGCGCGCG	313		
Qy	242	GCAGAACGGGC----AGGTGTCCAGATTGCGGCGAGCGAGCTGCGCCTTGCCCCGGC	296		
Db	314	GCGCGCGGGCGGAGGCGCGCGTGCAGATATGCGGCGAGCGAGTGGGGTGGCTTGAC	373		
Qy	297	GGGACCCCTTCGTGGCGGTGCAACGAGTGCCTTCCCCGTCTGCGCGGACTGTCTACGA	356		
Db	374	GGGAGCCCTTCGTGGCGGTGCAACGAGTGCCTTCCCCGTCTGCGCGCGCTGTACGAG	433		
Qy	357	TACGAGCGCGGGAGGGGACGCGAAGTGCCTCCAGTGAAGACTCGATACAGCGCCTC	416		
Db	434	TACGAGCGCGGGAGGGGCTCGCAAGCGTGCCTCCAGTGAAGACTCGATACAGCGCCTC	493		
Qy	417	AAGGCTGCGCAACGTGTGACCGGTGACGAGGAGGAGGACGCGCGTGTGATGACCTGACAAC	476		
Db	494	AAGGCTGCGCGCGGGTGGCGCGGACGAGGAGGAGGACGCGCGTGTGACGACCTGAGGGGC	553		
Qy	477	GAGTTCAACTG-----GACGGCATGACTCGCAGTCTGTGGCGCGAGTCC	521		
Db	554	GAGTTCGGCTGACGAGACGCGCGCGCCCAAGAGACGACCGCAGTACGTGCGCGAGTCC	613		
Qy	522	ATGCTCTACGGCCACATGAGCTACGCGGTGAGGTGACCTTAATGGCGCGCACAACT	581		
Db	614	ATGCTCAGGGCGGAGATGAGCTACGCGCGCGCGCGCA-----CGCGCACCCCGGC	664		
Qy	582	TTCCAGCTCAACCCCAATGTTCCACTCTCAACCAACGGGCAATGTGTGATGATCCCA	641		
Db	665	TTCAAGCCCCCTCCCAACGTGCGCGCTCTCAACCAACGGGCAATGTGTGATGATCCCG	724		
Qy	642	CCGAGACGACGCGCGCTGTGCTTCTTTCATG-----GGTGTGGGGGAAAGAG	692		
Db	725	CCGAGACGACGCGCGCTGTGCTTCTTTCATGAGCGCGCGCGCGCGCGGCGCAAGAG	784		
Qy	693	ATACATCCCTTCCTTATGCGGATCCAGCTTACCTGTGCAACCCAGSTCTATGACCCA	752		
Db	785	ATCACCCGCTCCCTTTCGAGATCCCACTTCCAGTGAACCGAGATCCATGGAACCG	844		
Qy	753	TCCAAGATCTTGCTGCATATGGGTATGTGTGCTTGAAGCAACGATGAGAAAT	812		
Db	845	TCCAAGATCTTGCGCGCTACGATATGGCAGGTGGCTGGAAGAGAGATGAGGGC	904		
Qy	813	TGGAAGCAAGACAAGAGAGATGACCAACGGGGAATGATGTGTGTGAT--GAT	869		
Db	905	TGGAAGCAAGACAAGAGAGCGCCTGCAAGATGTCAAGAGCGAGGTGCGGTGATTTGGGAT	964		
Qy	870	GGTACGATGCTGATCTACCACTAATGATGAAGCAAGACAACACTGTCCAGAAAATT	929		
Db	965	GGGACGATGAGATCTGCACCTAATGATGAAGCTAGGCAAGCAATGTCCAGAAAAGTTC	1024		
Qy	930	CCACTTCATCAAGCCAGATTAATCCATATAGGATGATTATCATTAATTCGGCTGTGTT	989		
Db	1025	CCTATATCAACAACCGAATTAATCCCTACAGAGATGATTATCGTTATCCGGTGTGTT	1084		
Qy	990	TTGGGTTCTTCTTCCACTACCGAGTGTGATCCGGTGAATGATGATTTGCTTGTGG	1049		
Db	1085	TTGGGTTCTTCTTCCACTACCGAGTGTGATCCGGCGAAAAGATGATTTGCAATGTGG	1144		
Qy	1050	CTCATATCTGTATCTGTGAATCTGTTGCCATGTCTTGATTTTGATCAATTCCCA	1109		
Db	1145	CTCATATCTGTATCTGTGAATCTGTTGGCATGTCTTGATTTTGATCAAGTTCCCA	1204		
Qy	1110	AAGTGTCCCTATTGAGAGAGACTTAACCGGCTGTCACTGAGCTTGACAAG	1169		
Db	1205	AAGTGTCCCAATCGAAGAGAGACTTAACCTGACCGTTGTCACTAAGGTTTGACAAG	1264		

Qy	1170	GAAAGCCAGCCATCTCAACTTGTCTCCAATTGATTTCTTTGTCAGTACGGTTGATCCCTTA	1229		
Db	1265	GAAAGTCAACCTCTCAGCTTGTCTCAATCGACTTCTTTGTCAGTACGGTTGATCCACA	1324		
Qy	1230	AAGAACTCTTTTGTCAACAATACTGTCTATCTATCTTTCGGTGAATATCTT	1289		
Db	1325	AAGAACTCTCTTGTGTACAGCAACACTGTCTCTTCCATCTTCTGTGATTAATCCG	1384		
Qy	1290	GTTGATTAAGTTCTTGTCTATGTTTCTGATGATGTGTGCAATGCTAACGTTGAACA	1349		
Db	1385	GTTGATTAAGTTCTTGTCTATGTTTCTGATGATGTGTGCAATGCTAACGTTGAACA	1444		
Qy	1350	TTATCTGAACATCTGAATTTGCAAGAATGGGTTCTTCTGCAACGGTACATATT	1409		
Db	1445	TTGTCTGAACATCTGAATTTGCAAGAATGGGTTCTTCTGCAAAAAGTTTAATATC	1504		
Qy	1410	GAACCTCGCGCTCCAGAGTGTACTTCCAAACAGAAATAGACTACTTGAAAGACAAGTG	1469		
Db	1505	GAGCTCTGTCTCTGAGTGTACTTCCAAACAGAAATAGACTACTTGAAAGACAAGTT	1564		
Qy	1470	GCAACAACCTTTGTTAGGAGAGAGAGCAATGAAGAGAGATGAGGAATTCAGGTG	1529		
Db	1565	GCTGCTTCATTTGTTAGGAGAGAGAGCGCATGAAGAGAAATACGAGGAATTCAGGTA	1624		
Qy	1530	AGAATCAATGCCTTAGTTCCAAGCCAGAAAGTTCTGAAGAAGATGACAAATGCA	1589		
Db	1625	AGATCAATGCCTTAGTTCCAAGCCAGAAAGTTCTGAAGAAGATGACAAATGCA	1684		
Qy	1590	GATGAACCCCTGCGCTGGAACAATGTTCTGTATCATCTTGAATGATTCAGGTCTTC	1649		
Db	1685	GATGAAGCCCTGCGCTGGAACAACGTACCGCATCTTGAATGATTCAGGTATTC	1744		
Qy	1650	CTTGCCCAAGCGGAGGCTTGACTGTGAGGGAATGAACGCCAGATTGTTATGTT	1709		
Db	1745	CTTGCCCAAGTGGCGGTGTGATGTGAAGGAATGAGTTGCCCTGCTGTATATGTC	1804		
Qy	1710	TCTAGAGAAAAGCAGCAGGCTATTAACCATCATTAAGAAAGCTGTGCTATGAATGACTG	1769		
Db	1805	TCGAGAGAAAAGAGGCCAGGTATTAACCATCACAAAGAGCTGTGCCATGAATGACTG	1864		
Qy	1770	GTCGAGTCTCTGCTGTACTAACAATGCTCCATATTTGTTAACTTGATGTGATCAC	1829		
Db	1865	GTCGAGTCTCTGCTGTACTAACAATGCTGCATACCTATGAACTTGACTGTGATCAC	1924		
Qy	1830	TACATCAACAACAGCAAGGCTATTAAGAGCAATGTGTTTATGATGACCCCTTACTA	1889		
Db	1925	TACATCAACAATAGCAAGGCCATTAAGAGGCTATGTGTTTCAATGATGATCTTTGGTG	1984		
Qy	1890	GGAAAGAGGTTTGCTATGTACAGTTCCTCAAAAGATTGATGGATGTATGCCATGAC	1949		
Db	1985	GGGAAGAAAGTGTCTATGTACAGTTCCTCAGAGGTTGATGTATTGCAAAAATGAT	2044		
Qy	1950	CGATATGCTAACCGGAATGTGCTTTTGTGATATCAACATGAAGGTTTGATGTGAT	2009		
Db	2045	CGATACGCTAACAGGAAGTGTCTTTTGTGACATCAACATGAAGGTTTGACGCTATT	2104		
Qy	2010	CAGGTCCAATTATGTGTGTACTGATGTGATTTAGAGCAGGCATTATATGTTAT	2069		
Db	2105	CAAGAACCCATTATGTGTGTACTGATGTGTTTCAAGCGGCAAGCACTGTATGTTAT	2164		
Qy	2070	GATGCCCCCAAAACAAAGAGCCACATCAAGGACTTGCAACTGCTGGCCCAAGTGTGC	2129		
Db	2165	GATGCTCTTAAACGAAGAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGTGC	2224		
Qy	2130	TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAACTAACCAACCCAAAACAGAG	2189		
Db	2225	CTCTCTGCTGCTGACAGAAACAAGATTAAGAAAGAACTAACCAAAACCAAGAGCAG	2284		
Qy	2190	AAGAAAAAGTTATTATTTTCAAGAAAGAGAAACCAATCCCTGCATATGCTCTTGGT	2249		
Db	2285	AAGAAAGAAAGATTATTTTTCAAGAAAGCAAGAAACCAATCTCTGCATATGCTTTGGT	2344		

OY		2250	GAAATTGACGAAGCTGCTCCAGGAGCTGAGAATGAAAAGGCCGGTATTGTAAATCAACAA	2309
Db		2345	GAAATTGATGAAGTGCTCCAGGTGCTGATATCGAAGAGGCCGGAATCGTAATCAACAG	2404
OY		2310	AAATTGAAAAAGAAATTGGCCAATCTTCTGTTTTTTGTACATCCACACTTCGAGAAT	2369
Db		2405	AAACTAGAGAAGAAATTGGGCAGTCTTCTGTTTTGTCCGATCAACAACCTTGTGAGAAC	2464
OY		2370	GGTGAACCTTGAAGAGTGCAGTAGTCCCTGCTTCTCTTTGAAAAGAGCTATACATGTCATT	2429
Db		2465	GGAAGGACCCTGAAGAGCGCAAGTCCAGCTTCTCTCTGAAGGAAGCTATACATGTATC	2524
OY		2430	AGTTGTGTTATGAAGACAAGACAAGACTGGGGAAAAAGATTGGCTGGATCTATGCATCA	2489
Db		2525	AGCTGCGGCTACGAAGACAGAACCGACTGGGGAAAAAGATTGGCTGGATTTACCGATCG	2584
OY		2490	GTTACAGAAGATATTCTAACTGGTTTCAAGATGCATTGTGTTGGCGGTCAAATTTAC	2549
Db		2585	ATCACAGAGATATCTGACTGGAATTTAAGATGCACTGCCATGGCTGGCGGTCTATTAC	2644
OY		2550	TGCATACCTAAACGGGTTCGATTCAAAGGTTCTGCACCTCGAATCTTTCAGATCGTCTT	2609
Db		2645	TGCATCCCCGAAGCGCGCTGCATTCAAAGGTTCTGCGCTCTGAACCTTCCGACCGTCTT	2704
OY		2610	CACCAGTGCTTCGGTGGGCTCTTGGGTCTATGTAGATCTTCTTCAGCAATCATGCCCC	2669
Db		2705	CACCAGGTCTTCGCTGGGCGCTTGGGTCCGTCGAATTTCTTCAGCAAGCACTGCCCA	2764
OY		2670	CTTTGGTATGGGTATGTTGGCGGCTCGAATTTTGTGAAGAATTTTCTCATCAACTCC	2729
Db		2765	CTTTGGTACGGATACGGCGCGCGGCTAAATTTCTTGGAAGGTTTCTTATATCACTCC	2824
OY		2730	ATCGTGATCCTTGGACATCTATTTCCCCTCTTGCGTTACTGTACATTGCGTCCCATCTGT	2789
Db		2825	ATCGTTTATCCCTGGACGTTCATTCTCTCTGCGTTACTGTACCTTGCTGCCATCTGC	2884
OY		2790	TTATTGACAGGAAATTATCACTCCAGAGCTGAATAAATGTTGCCAGCCTGTGTTCATG	2849
Db		2885	CTGCTCAGCGGAAGTTTATCACACAGAGCTTACCAATGTCGCCAGTATCTGTTTCATG	2944
OY		2850	TCACCTTTTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGAGTGTGTTGA	2909
Db		2945	GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGMAATGAGGTGAGTGGCGTGGCC	3004
OY		2910	ATTGATGACTGTGAGGAATGACAGTTCCTGGGTCAATGGAGGTGTCTCTCAACCTC	2969
Db		3005	ATCGACGACTGTGTGAGGAACGACAGAGTTCCTGGGTCAATCGAAGCGCTTTCGCGCATCTG	3064
OY		2970	TTTGTCTGTCTCCAGGAGCTTCTCAAGTCAATAGCTGTGTGTTGATACAAAGCTTCACCGTG	3029
Db		3065	TTGCGGCTGTTCCAGGGCGCTGCTGAAGGTGTTGCGCGGCATCGAACAGACTTCACCGTG	3124
OY		3030	ACATCAAGGGGTGAGATGATGAGAGTTCCTCAGAGCTATATACTTCAATGGAATGACTAAC	3089
Db		3125	ACGTTCGAAGCGCGGGAGCGACGAGAGGTTCTCGAGCTGTACACGTTCAAGTGGACCAAC	3184
OY		3090	TTATTGATACCTCCTACCACTTGCTTCTATTGAACTTCATGTGTGTGCTGCTGGCGTT	3149
Db		3185	CTGCTGATACCCCGACCAAGCTTCCTCTGCTGAACCTTCATCGGGGTGTGCGCGGATC	3244
OY		3150	TCAAATGCGATCAATAACGGAATGAGTCATGGGCGCCCTCTTTGGGAAGCTATTCTTT	3209
Db		3245	TGGAACCGCATCAACAACGGGTACAGAGTCTGTGGGCGCCCTGTTCGGGAAGCTCTTCTTC	3304
OY		3210	GCAATTTGGGTGATGTGCCATCTTTATCCCTTTCTCAAAGGTTTGTGTGGAAGCAAAAC	3269
Db		3305	GCCTTCTGGGTGATCGTCCACTGTATACCCGTTCTCAAGGTTCTGTGGGAGGCAAGAAC	3364
OY		3270	AGGACACCACGATGTGTCATCGTCTGCTCCATTCCTGTGCGTTCAATCTTCTCGCTCCTT	3329
Db		3365	AGGACCGCGACGATCGTCATCGTCTGCTCCATCTGCTGCGCTCGATCTTCTCGCTCCTG	3424
OY		3330	TGGGTTGGAATGATCTCTTCTCTTCCGGAAGATGATGTCGCGTCTCTTGAGAGATGTGCT	3389

Db	3425	TGGGTCCGCGTCGACCCGTTCTCTCGCCAGAGCAACGCCCGCTCTGAGAGTGTGC	3484
Qy	3390	TTGATTGCAACT	3402
Db	3485	CTGACTGCAACT	3497
RESULT 8			
AAZ99524	ID	AAZ99524	standard; DNA; 3813 BP.
XX	AC	AAZ99524;	
XX	DT	03-JUL-2000	(first entry)
XX	DE	DNA encoding a maize cellulose synthase.	
KM	Maize;	cellulose synthase; stalk quality; stand; silage; cellulose;	
KM	transgenic	plant; plant breeding marker; ss.	
XX	Zea	mays.	
XX	OS		
XX	Key	Location/Qualifiers	
FT	CDS	215..3493	
FT		/*tag= a	
FT		/product= "cellulose synthase"	
FT		/note= "no termination codon given"	
XX	XX	WO200009706-A2.	
XX	XX	24-FEB-2000.	
XX	PF	16-AUG-1999; 99WO-US18760.	
XX	17-AUG-1998;	98US-0096822.	
XX	(PION-)	PIONEER HI-BRED INT INC.	
PI	Dhugga	KS,	Helentjaris TG, Bowen BA, Wang X;
XX	XX		
DR	WPI;	2000-224343/19.	
DR	P-PSDB;	AAV84118.	
XX	XX	New genes which encode maize cellulose synthase polypeptides in plants	
PT	useful	for modulating the expression of cellulose synthase in plants	
PT	and to	produce transgenic plants expressing the novel protein -	
PS	Claim 1;	Page 168-173; 119pp; English.	
XX	XX	The present sequence encodes a maize cellulose synthase polypeptide.	
CC	CC	The cellulose synthase can be used for the improvement of stalk quality	
CC	CC	for improved stand or silage. It also provides an increased concentration	
CC	CC	of cellulose in the pericarp, hardening the kernel and improving its	
CC	CC	handling ability. The sequences are used to produce transgenic plants	
CC	CC	and seeds expressing the cellulose synthase. The polynucleotide is	
CC	CC	used for modulating, preferably increasing, the level of the synthase	
CC	CC	in a plant cell. The plants are preferably monocots. The polynucleotide	
CC	CC	is also used as a probe or primer in the detection quantitation or	
CC	CC	isolation of gene transcripts. The probes are useful in detecting	
CC	CC	deficiencies in the level of mRNA in screenings for desired transgenic	
CC	CC	plant, for detecting mutations in the gene, for monitoring upregulation	
CC	CC	of expression or changes in enzyme activity in screening assays of	
CC	CC	compounds, for detection of any number of allelic variants of the gene,	
CC	CC	or for use as molecular markers in plant breeding programs. The	
CC	CC	isolated nucleic acids of the present invention can also be used as	
CC	CC	recombinant expression of their encoded polypeptides or for use as	
CC	CC	immunogens in the preparation and/or screening of antibodies. The	
CC	CC	proteins can be employed in assays for enzyme agonists or antagonists	
CC	CC	of enzyme function or for use of immunogens or antigens to obtain	
CC	CC	antibodies specifically immunoreactive with a protein.	
XX	XX	Sequence 3813 BP; 873 A; 967 C; 1111 G; 862 T; 0 other;	



Query Match	56.9%;	Score 2241;	DB 21;	Length 3813;
Best Local Similarity	81.2%;	Pred. No. 0;		
Matches 2691;	Conservative	0;	Mismatches 580;	Indels 42;
				Gaps 6;
QY	123	GGAGTGAAGGGAGGAAGCGATGGAGCGCGCGGCTGTGGCCGCTCCCAAC	182	
Db	194	GGCGGGGGGGGGCTGAGATGAGGCTAGCGGGGGCTGTGGCCGCTCGATAAC	253	
QY	183	CGCAACGAGCTCGTCGTCATCCCGCCGACGGCGA-TCCCGGGCCGAAGCCCGCGCGA	241	
Db	254	CGGAACGAGCTGTGTGTGATCCGCCCGCACCCCGAGTCGGGAGCCCGGGCGCGCG	313	
QY	242	GCAGAACGGGC-----AGGTGTGCCAGATTGGCGGGAAGACGTGGCCTTGCCTCCGCG	296	
Db	314	GCGCGCCGGGCGAGCGCCGCTGCCAGATATGCGGCAAGAGTCGGGTGGCTTCGAC	373	
QY	297	GGGAGCCCTTCGTGGCGGTGCAACGAGTGGCGCTTCCCGTCTGCCGGGACTGTACGA	356	
Db	374	GGGAGCCCTTCGTGGCGGTGCAACGAGTGGCGCTTCCCGTCTGCCGGCTGTACGAG	433	
QY	357	TACGAGCCCGGAGGGGCAACGACACTGCCCCAGTGCAGACTCGATACAGCGCCTC	416	
Db	434	TACGAGCCCGGAGGGGCTCCGAAGCGTGCCTCCGACGTGCAAGACCCGCTACAGCGCCTC	493	
QY	417	AAGGGCTGCCAACGTGTGACCGGCTGACGAGAGGAGGACGGCGTGCATGACTTGAACAAC	476	
Db	494	AAGGCTGCCCGGGGTGGCCGGCGACGAGAGGAGGACGGCGTGCACGACTTGAAGGGC	553	
QY	477	GAGTTCAACTG-----GACGGCCATGACTGCGAGTCTGTGGCGAGTCC	521	
Db	554	GAGTTCGAGCTTGACAGACGGCGCCGCCACGAGAGACCCGACGATGCGCGAGTCC	613	
QY	522	ATGCTCTACGGCCACATGAGCTACGGCCGTGAGGTGACCTTAATGCGCGCCACAAGCT	581	
Db	614	ATGCTCAGGGCGCGACATGAGCTACGGCCGCGCGCGCA-----CGCGCACCCCGGC	664	
QY	582	TTCAGCTCAACCCCAATGTTCCACTCCTCAACCAAGGCCAATGTGTGATGACATCCCA	641	
Db	665	TTCAGCCCGCTCCCAACGTGCCGCTCCTCAACCAAGGCCACAGATGTTGATGACATCCCG	724	
QY	642	CCGAGCAGCACGCGCTGTGCTCTTTCATG-----GGTGTGGGGAAAGAGG	692	
Db	725	CCGAGCAGCACGCGCTGTGCTCTCAATGACGGCGCGCGGGGCAAGAGG	784	
QY	693	ATACATCCCCCTTCTTATGCGGATCCAGCTTACTGTGCAACCCAGTCTATGACCCA	752	
Db	785	ATCACCCGCTCCCTTTCGAGATCCCAACCTTCCAAGTCAACCGAGATCCATGACCCG	844	
QY	753	TCCAAGATCTTGTGTCATATGGGTATGTAAGTGTGTTGTTGGAAAGAACGGATGAGAAAT	812	
Db	845	TCCAAGATCTGCGCGCTTACGGATATGGCAGCGTGGCTTGAAGAGAAATGAGGGC	904	
QY	813	TGGAAGCAGAGACAAGAGGATGCACCAGACGGGGAATGATGTTGTGTGAT--GAT	869	
Db	905	TGGAAGCAGAGCAGGAGCGCTGCAGCATGTCAAGACGAGGGTGGCGGTGATGGGAT	964	
QY	870	GGTGACGATGCTGATCTACCACTAATGATGAAGACAGACAACACTGTCCAGGAAATTT	929	
Db	965	GGGACGATGCAGATCTGCCACTAATGATGAAGCTAGGACAGCCATTGTCCAGAAAGTTC	1024	
QY	930	CCACTTCCATCAAGCCAGATTAAATCCATATAGATGATTAATCATTAATTCGGCTGTGTT	989	
Db	1025	CCTATATCATCAAGCCGAATTAATCCCTACAGATGATTAATCGTTATCCGGTGTGTT	1084	
QY	990	TTGGGGTCTTCTTCCAATAACCGAGTGATGATCCCGTGAATGATGATTTGCTTTGTGG	1049	
Db	1085	TTGGGTTCTTCTTCCAATAACCGAGTGATGATCCCGGCAAGATGCAATTTGCATTGTGG	1144	
QY	1050	CTCATATCTGTTATCTGTGAATCTGTTGGCCATGTTTGGATTCTTGATCAATTTCCA	1109	
Db	1145	CTCATATCTGTAATCTGTGAATCTGTTGGATGTCCTGGAATTTTGATCAGTTCCCA	1204	

QY	1110	AAGTGGTTCCTTAATTGAGAGAGACTTAACCTAGACCGGGCTGCACTGAGGTTGCAAG	1169
Db	1205	AAGTGGCTTCCAATCGAGAGAGACTTACCTGACCGTTTGTCACTAAGGTTTGACAAG	1264
QY	1170	GAAGGCCAGCCACTCTCAACTTGCTCCAAATTGATTCTTGTGCTAGTACGGTTGATCCCTTA	1229
Db	1265	GAAGGTCAACCCCTCTCAGCTTGCTCCAATGCACCTCTTGTGCTAGTACGGTTGATCCACA	1324
QY	1230	AAGGAACCTCCTTTGGTCACACAATACTGTTCTATCTATCCTCTTCGGTGAATTACCT	1289
Db	1325	AAGGAACCTCCCTTGCTGTCACAGCGAACACTGTCCTTCCATCCTTCTGTGATTAATCCG	1384
QY	1290	GTTGATAAGGTTTCTTGCTATGTTTCTGATGATGAGTGCTGCAATGCTAACGTTGAAGCA	1349
Db	1385	GTTGAGAAGGTCCTCTGCTATGTTCTGATGATGAGTGCTGCAATGCTTAACGTTGAAGCA	1444
QY	1350	TTATCTGAACAATCTGAATTTGCAAGAATGGGTTCTTTCTGCAAAACGGTACAAATATT	1409
Db	1445	TTGTCGTAACAATCTGAATTTGCAAGAATGGGTTCTTTGACGAAAAAGTTTAATATC	1504
QY	1410	GAACCTCGCGCTCCAGAGTGCTACTTCCACACAGAATAGACTACTTGAAGAACAAGGTG	1469
Db	1505	GAGCCTCGTCTCTGAGTGTACTTCCACAGAGATAGACTAACCCTGAAGAACAAGTT	1564
QY	1470	GCAGCAAACTTGTGTAGGAGAGAGAGCAATGAAGAAGATATGAGGAATTCAGGTG	1529
Db	1565	GCTGCTTCATTTGTTAGGAGAGAGGCGATGAGAAGAATACGAGGAATTCAGGTA	1624
QY	1530	AGAATCAATGCCTTAGTTGTCCAAAGCCCAAGATTCCTGAAGAAGATGACAAATGCA	1589
Db	1625	AGGATCAATGCCTTGTTGTCAAAAGCCCAAAAGTTCTCGAAGAAGATGACAAATGCA	1684
QY	1590	GATGGAACCCCTGGCCTTGAAACAATGTTGATCATCCTGGAATGATTCAAGTCTTC	1649
Db	1685	GATGGAAGCCCTGGCCTTGAAACAACGTACGATCATCCTGGAATGATTCAAGTATTC	1744
QY	1650	CTTGCCAAAAGCGGAGGCTTGACTGTGAGGAAATGAACCTGCCACGATTGGTTATGTT	1709
Db	1745	CTTGCCAAAAGTGGCGGTGCTGATGTGAAGAAATGAGTTGCCCTGGCTGTTATGTC	1804
QY	1710	TCTAGAGAAAAGCAGCAGGCTATTAACCATCATTAAGAAAGCTGGTCTATGAATGCATTG	1769
Db	1805	TGAGAGAAAAGAGGCCAGGTTATTAACCATCACAAGAAGCTGGTCCATGAATGCATG	1864
QY	1770	GTCCGAGTCTTGCTGTACTAACCAATGCTCCATATTTGTTAACTTGATTGTGATCAC	1829
Db	1865	GTCCGTGCTCTGCTGTCTTATCAAAATGCTGCAATCTATTGAACCTTGACTGTGATCAC	1924
QY	1830	TACATCAACAACAGCAAGGCTATTAAGGAAGCAATGTGTTTATGATGAGCCCTTACTA	1889
Db	1925	TACATCAACAATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGATCCTTTGGTG	1984
QY	1890	GGAAGAAGGTTTGCTATGTACAGTTCCTCAAGAATTTGATGGATTGATCGCCATGAC	1949
Db	1985	GGAAGAAGAGTGTCTATGTACAGTTCCTCAAGGTTTGATGGTATTGACAAAATGAT	2044
QY	1950	CGATATGCTAACCGGAATGTGCTTTTGTGATATCAACATGAAGGTTTGATGGTATT	2009
Db	2045	CGATACGCTAACAGAAACGTGCTTTTGTGACATCAACATGAAGGTTTGACGGTATT	2104
QY	2010	CAGGTCCAATTTATGTTGGTACTGATGTGTAATTGAAGGACAGCATTATATGTTAT	2069
Db	2105	CAAGGACCCCAATTTATGTGGTACTGATGTGTTTCAAGCGGACACTGTATGGTATT	2164
QY	2070	GATGCCCCCAAAACAAGAAGCCACCATCAAGACTTGCAACTGCTGGCCCAAGTGCTC	2129
Db	2165	GATGCTCCTAAACAAGAAGCCACCATCAAGACTTGCAACTGCTGGCCCAAGTGCTC	2224
QY	2130	TTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAAAACCAAAACAGAG	2189
Db	2225	CTCTCTGCTGCTGCAAGCAAGAACAGAAATTAATAAGAAAGCTACAAAACCAAAAGCGAG	2284
QY	2190	AAGAAAAAGTTATATTTTTTCAGAAAGAAGAACCAATCCCTGCATATGCTCTTGCT	2249

Db 2285 AAGAGAAAAGATTATTTTTCAGAAAGCAGAAAAACCCATCTCTGCAATATGCTTGGGT 2344  
QY 2250 GAAATTGACGAAGCTGCTCCAGAGCTGAGATGAAAAAGCCGGTATTGTAATCAACAA 2309  
Db 2345 GAAATTGATGAGGTGCTCCAGGTGCTGATATCGAAGAGCCGGAATCGTAATCAACAG 2404  
QY 2310 AAATTAGAAAAGAAATTTGGCCAATCTTCTGTTTGTACATCCACTTCTCGAGAAAT 2369  
Db 2405 AAAC TAGAGAGAAATTTGGGAGTCTTCTGTTTGTGTCATCAACACTTCTTGAGAAC 2464  
QY 2370 GGTGAACCTTGAAGAGTGCAGAGTCCGTCTCTTTTGAAGAAGCTATACATGTCATT 2429  
Db 2465 GGAGGAGACCTGAGAGAGCGCAAGTCCAGCTTCTTCTGAAGAGAGCTATACATGTTATC 2524  
QY 2430 AGTGTGTTATGAACAACAAGACAGACTGGGAAAAGAGATTGGCTGATCTATGATCA 2489  
Db 2525 AGCTGGGCTACGAAAGACAAAGACCGACTGGGAAAAGAGATTGGCTGATTTACGATCG 2584  
QY 2490 GTTACAGAGATATTTCTAACTGTTTCAAGATGATGTCATGTTGGCGGTCAATTAC 2549  
Db 2585 ATCAGAGAGATATCTGACTGATTAGATGCACTGCCATGCTGGCGGTCTATTAC 2644  
QY 2550 TGCATACCTTAAAGGGTTCATTTCAAAAGTTTCGACCTCTGAATCTTTCAGATCGTCTT 2609  
Db 2645 TGCATCCCGAAGCGGCTGCATTTCAAAGTTTCGCGCTCTGAACCTTTCGACCGTCTT 2704  
QY 2610 CACCAGTGTGGTGGGCTCTGGGCTCTGATGATGATCTTCTTCAAGATCATTTGCCCT 2669  
Db 2705 CACCAGTCTGCTGGGCGCTTGGGCTCGGTGAATTTCTTTCAGAGACACTGCCCA 2764  
QY 2670 CTTGGTATGGGTATGTTGGCGGCTGAAATTTTGGAAAGATTTTCTTACATCAACTCC 2729  
Db 2765 CTTTGTACGGATACGGCGGGGCTAAATTTCTTGAAAGGTTTCTTATATCAACTCC 2824  
QY 2730 ATCGTATCTTGGACATCTATTTCCCTCTTGGCTTACTGTACATTTGCCCTCATCTGT 2789  
Db 2825 ATGTTATCTTGACGTCCATCTCTCTGCTTACTGTACTGTACTGTGCTGCACTGCTC 2884  
QY 2790 TTATTGACAGGAAATTATCACTCCAGAGCTGAATATGTTGCCAGCTGTGTTCATG 2849  
Db 2885 CTGCTCAGCGGAAATTATCACACAGAGCTTAACTATCGCCAGATCTGTTCTATG 2944  
QY 2850 TCACCTTTTATCTGCAATTTTGTCTACGAGCACTTGAATGAGATGAGTGTGTTGA 2909  
Db 2945 GCACTTTTCACTGATCTCCGTGACCGGCACTCGAATAGAGTGAAGTGGCGTGGCC 3004  
QY 2910 ATTGATGACTGTGAGGAATGAGCAGTTCTGGTCAATTGGAGTGTGCTCACACCTC 2969  
Db 3005 ATCGACGACTGTGAGGAAGCAGCAGAGTCTGGGTATCGGAGCGCTTTCGCCCATCTG 3064  
QY 2970 TTTGCTGTGTTCCAGGGAATCTTCAAGTCAATAGCTGTGTTGATACAACTTCAACGTG 3029  
Db 3065 TTGCGGTGTTCCAGGGCTGCTGAAAGTGTGCGCGGCATCGACACGAGCTTCAACGTG 3124  
QY 3030 ACATCAAAAGGTTGAGATGATGAGAGTTCTCAGAGCTATATACATCAATGAAGTACC 3089  
Db 3125 ACGTGAAGGCCGGGAGCAGCAGAGTCTCGAGCTGTACACGTTCAAGTGAACACC 3184  
QY 3090 TTATTGATACCTCTTACCACCTTGTCTTATGAACCTTCAATTGTTGTTGCTGCGCTT 3149  
Db 3185 CTGCTGATACCCCGACCAAGCTCTCTGCTGAACCTTCACTGGGTGTGGCGGATC 3244  
QY 3150 TCAATGCGATCAATAACGATATGATGATGAGGCCCCCTCTTTGGAGACTATTCTTT 3209  
Db 3245 TCGAAGCGATCAACAACGGGTAGAGTGTGGGGCCCCCTGTTGGGAAGCTCTTCTTC 3304  
QY 3210 GCATTTGGGTGATGTCCATCTTATCCCTTCTCAAAGTTTGGTTGGAAGGCAAAAC 3269  
Db 3305 GCCTTCTGGGTGATGCTCCACCTGTACCCGTTCTCAAAGGTTCTGGTGGGAGGCAAGAC 3364  
QY 3270 AGGACACCAACGATGTGATGCTGTGTTCAATTTGCTGCTTCTTCTGCTCCTT 3329

Db 3365 AGGACGCCGACGATCGTATGCTGTGCTCATCTCTGCTGAGCTTGATCTTCTGCTCTG 3424  
QY 3330 TGGGTTCCGATTGATCTTCTTCTGCGAAGATGATGTCGGCTTCTTGAGAGTGTGT 3389  
Db 3425 TGGGTCCGCGTCGACCCGTTCTCTCGCCAAGACGAGCGCCGCTCTGAGAGTGTGCG 3484  
QY 3390 TTGATTTGCAACT 3402  
Db 3485 CTGACTGCAACT 3497

RESULT 9  
ABZ14725  
ID ABZ14725 standard; DNA; 3255 BP.  
XX  
AC ABZ14725;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2530.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US26685.  
XX  
PR 24-AUG-2000; 2000US-227866P.  
PR 26-JAN-2001; 2001US-264647P.  
PR 22-JUN-2001; 2001US-300111P.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
DR WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed  
PT and producing plants with increased tolerance to these abiotic stresses  
PT  
XX  
PS Claim 144; SEQ ID NO 2530; 577bp + Sequence Listing; English.  
XX  
CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising:  
CC (a) contacting nucleic acid representative of expressed polynucleotides  
CC in the plant cell with an array or probes representative of the plant  
CC cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.  
XX  
SQ Sequence 3255 BP; 841 A; 641 C; 813 G; 960 T; 0 other;  
Query Match 38.3%; Score 1509; DB 24; Length 3255;  
Best Local Similarity 67.6%; Pred. No. 0;  
Matches 2207; Conservative 0; Mismatches 1030; Indels 30; Gaps 5;  
QY 144 ATGAGCGGAGCGCGGCTGTGCGCGGCTCCCAACGCGCAAGAGCTGCTGTCATC 203  
Db 1 ATGAACACCGGTGTGCGGTATATCGCCGTTCTCACAAAGAAATGAGTTGTCTCATTT 60

QY 204 CGCCGCGACGGCGATCCCGGGCCGGAAGCCCGCGGAGACGAACGGGCGAGGTGCGCAG 263  
Db 61 AATGCCGATGAAATGCCGAATAGATCAGTCCAAAGAGCTGAGTGACAGACATGTCAA 120  
QY 264 ATTGCGGCGACGACGTCGGCCTTGCCCCCGGCGGGGAGCCCTTCGTGGCGTGCAAG 323  
Db 121 ATTCGACAGATGAGATCGAATGACTGTGTGAGGAACCGTTGTGGCATGTAAAGAA 180  
QY 324 TGGCCTTCCCGCTCTGCGCGGAGCTGCTACGAATACGAGCGCGGAGGCGACGCAAGAC 383  
Db 181 TGTGCATTCCTGTGTAGACCTTGCTATAGATACGAAGACGAGAAGGCAATCAGCT 240  
QY 384 TGCCTCCAGTGCAGACTCGATACAGCGCCTCAAGGGCTGCCAAGTGTGACCGGTGAC 443  
Db 241 TGTCCACAGTGCAGAAACCCGTTCAACGTTCTTAAAGGAGTCCAAGAGTTGAAGGTGAT 300  
QY 444 GAGGAGGAGGACGCGCGTCGATGACTGGACAACGAGTTCAACTGGGACGGCCATGACTCG 503  
Db 301 GAAGAGGAAAGATGACATTGATGATTTAGACAATGAGTTTGA--GTATGMAATTAATGGG 357  
QY 504 CAGTCTGCGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGAGGTGACCT 563  
Db 358 ATGGATTTGATCAGGTTTCTGAAGGTATGTCAATCTCTGTCGCAACTCCGGTTTCCCA 417  
QY 564 AATGCGCGCCACAGCTTTCCAGCTCAACCCCAATGTTCCACTCTCACCACGCGCAA 623  
Db 418 CAATCTGATTTGGATTGAGTCCAGCTCCAGCTGCGCTCAGATTCATGTGCTGACTTACGCGAC 477  
QY 624 ATGCTGATGACATCCACCGGACGACGACGCGC--TGGTCCCTTCTTTCATGGTGT 680  
Db 478 GAGGACGTTGAGATTTCTTCTGTATAGACATGCTCTTATTTGTTCTCTTCACTTGGTGT 537  
QY 681 GGGGAAAGAGGATACATCCCTTCTTATGCGGATCCCACTTACCTGTGCAACCCAGG 740  
Db 538 CATGGCAATAGAGTTCATCTGTGTTCTTCTTCTGACCCGACGTCGTCACATCCAGG 597  
QY 741 TCTATGGAACCCATCCAAAGATCTGCTCATATGGTATGGTATGTTGCTTGGAAGAA 800  
Db 598 CCTATGTAACCTCAGAAAGATCTGCGGTTATGTTATGAAAGTGTGCTTGAAAGAT 657  
QY 801 CGGATGGAATTTGGAAGCAGAGCAAGAGAGATGACCAAGCGGGA-----AT 851  
Db 658 CGGATGGAAGAATGGAAGAAAGACGAATGAGAAACCTTCAGGTTGTTAGCATGAAGA 717  
QY 852 GATGCTGTGTGATGATGCTGAAGATGCTGATCTACCACTAATGATGAAGCAAGCAA 911  
Db 718 GATCTGATTTTGAAGATGGTATGATGCTGATTTTCCATGATGATGAGGAGGACAG 777  
QY 912 CAACTGTCCAGGAAAATCCACTTTCATCAAGCCAGATTAAATCCATATAGATGATTATC 971  
Db 778 CCAATTGTCTAGGAAGATACCAATCAATCGAGCAAGATAATCTTACCGGATGTTAAT 837  
QY 972 AATTATCGGCTGTGTTTGGGCTCTTCTTCCACTACGAGTGAATCCGGTGAAT 1031  
Db 838 GTGCTACGTCTGTGATTCCTGCTCTTCTTCACTACCGTATTCCTTCAACCCGTCAAA 897  
QY 1032 GATGCATTTGCTTGTGCTCATATCTGTATCTGTGAATCTGTTTGCCATGCTTGG 1091  
Db 898 GATGCATATGCTTGTGCTTATTTCTGTATATGTGAGATATGCTTGTTCATGG 957  
QY 1092 ATTCCTGATCAATTCCTCAAGTGTTCCTTATGAGAGAGACTTACCTAGACCGGCTG 1151  
Db 958 GTTCTTGATCAGTTCCTCAATGTGTAACCTATCGAGCGGAGAAACGTACTTGACCGACTC 1017  
QY 1152 TCACTGAGGTTTCGACAAGGAGCCAGCCATCTCAACTTGTCCAATTGATTTCTTTGTC 1211  
Db 1018 TCAATTAAGATATGAGAAAGAGGAAACCGTCGGGACTATCCCTGTGATGTAATTTGTT 1077  
QY 1212 AGTACGGTTGATCCCTTAAAGAACTCTCTTGTGTACAACAAAATACTGTTCTATCTATC 1271  
Db 1078 AGTACAGTGAATCCATTGAAGAGCCTCCGCTTATTAAGCAAAATACTGTCTGTCTAAT 1137  
QY 1272 CTTTCGGTGATATATCTGTGTAAGGTTTCTGTATGTTTCTGATGATGCTGTGCA 1331

Db 1138 CTTGCTGTGATTAATCTGTGCGATAAGGTTGCTTGTATACGTATCTGATGATGCTGCT 1197  
QY 1332 ATGCTAACGTTTGAAGCATTATCTGAAACATCTGAAATTTTGCAAGAAATGCGTCTTTC 1391  
Db 1198 ATGCTTACTTTCGAGCTCTTCTTGAGACCGCTGAATTCGCAAGGAAATGCGTCTTTC 1257  
QY 1392 TGCAAACGGTACATATTTGAAACCTCGCGCTCCAGAGTGTACTTCCACAGAAAGATGAC 1451  
Db 1258 TGCAGAAATATTTGATTTGAGCCTCGTGCTCCGAATGATTTCTGCAATTAATGAGAC 1317  
QY 1452 TACTGAAAGACAAGGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAATGAAGAGAG 1511  
Db 1318 TACTGAAAGATTAAGTTCATCCCGCATTTGTTAGGAGCGGCGAGCCATGAAGAGAT 1377  
QY 1512 TATGAGAATTCAGGTGAGATCAATGCTTAGTTGCCAAAGCCAGAAAGTTCTTGAA 1571  
Db 1378 TATGAAGAATTCAAAGTAAAGATCAATGCTTAGTAGCAACAGCACAGAAAGTCCAG 1437  
QY 1572 GAAAGATGACAATGCAAGATGGAACCCCTGCGCTGGAACAATGTTGATCATCT 1631  
Db 1438 GATGTTGACTATGCAAGACGCTACACCTTGCGCGTAAATAGTGTCCAGATCATCT 1497  
QY 1632 GGAATGATTCAGGTTCTTCTTGCCAAAGCGGAGGCTTGACTGTGAGGAAATGAAGCTG 1691  
Db 1498 GGCATGATTCAGGTTCTTCTTGAGAGTGACGCGTGTCTGTATGTGCAAAACAAGAGTTG 1557  
QY 1692 CCACGATTTGTTATGTTTCTAGAGAGAAAGCAAGGCTATTAACCATCAATGAAGAGCT 1751  
Db 1558 CCTCGATTAGTTACGTTTCTCGTAGAAGAGAACCCGATTTGATCAGCAATGAAGAGCT 1617  
QY 1752 GGTGCTATGAATGATGCTGTCGAGTCTGCTGTAATAACAATGCTCAATATTGTTA 1811  
Db 1618 GAGGCTATGAATTCCTGATACGAGTCTTGCGGTTCTATCAATGCTCCTTACTCTG 1677  
QY 1812 AACTTGATTTGATCACTACATCAACAACAGCAAGGCTATAAGAGCAATGTGTTT 1871  
Db 1678 AATGCTGATTTGATCACTACATCAACAATAGCAAAAGCTCTTAGAGAGCAATGTGTTT 1737  
QY 1872 ATGATGACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTTCCTCAAAAGATTGAT 1931  
Db 1738 ATGATGATTCCTCAGTCAAGAAAGAAATCTGTTATGTTCAAGTTCCCTCAAAAGTTGAT 1797  
QY 1932 GGGATGATCGCCATGACCGATATGCTTAACCGGAATGTTGCTTTTGTATCAACATG 1991  
Db 1798 GGGATGATAGGCAAGATGATACTCAAAATGCAATGTGTGTTCTTTGATATCAATATG 1857  
QY 1992 AAAGTTGATGATTAATTCAGGTCCAATTAATGTTGTAAGTGTGATGATGTAATTAAGAG 2051  
Db 1858 AAAGTTGATGAGGCTACAAGGCGCTATATACGTCGGTACAGGTTGTGTTTCAAGAGG 1917  
QY 2052 CAGGCATTAATGCTTATGATGCTCCCAAAACAAGAGCCACCATCAAGACTTGCAAC 2111  
Db 1918 CAAGCGCTTACCGAATTTGATGACACCGAAGAGAAAGAGGCCACGTAAAGACATGCAAT 1977  
QY 2112 TGCTGGCCCAAGTGTCTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGAAAGACT 2171  
Db 1978 TGCTGGCCAAATGTGTCTCTATGTTTGGTTCAAGAAAGATCGTAAAGCAAAAGACA 2037  
QY 2172 ACCAAACCCAAACAGAGAAAGAAAGTATTAATTTTCAAGAAAGAGAAACCAATCC 2231  
Db 2038 GTGCGTCGATTAAGAAAGAAAGAA-----TAGGAAGCGTCAAAAGCAGATC 2085  
QY 2232 CCTGATATGCTCTTGTGAATTTAGCAAGCTGCTCCAGAGCTGAGAAATGAAGAGCC 2291  
Db 2086 CAGCATTAAGAAATATCGAAGAGGCGCGCTCACTAAAGTTCTTAAGTAAAGACATCA 2145  
QY 2292 GGTATTTGAATCAACAAATTAAGAAAGAAATTTGGCCATCTTCTGTTTGTGTTACA 2351  
Db 2146 ACCGAGGCAATGCAAAATGAAGTTGAGAGAGAAATTTGGGCAGTCTCTGTATTTGTTGCA 2205  
QY 2352 TCCACACTTCTCGAGAAATGCTGGAACCTTGAAAGAGTCCAAAGTCTGCTTCTTTTGAAA 2411



Db 2206 TCTGCGGATGAGAAATGGTGGATGGCTAGAAACGCAAGCCCGGCTGTCTGCTTAA 2265  
QY 2412 GAAGCTATACATGTCATTAGTTGTGTTATGAGACAAGACAGACTGGGAAAAGAGATT 2471  
Db 2266 GAAGCCATCCAAAGTCATTAGTTGCCGATATGAAATAAACTGAATGGGAAAAAGAGATT 2325  
QY 2472 GGCTGATCTATGATCAGTTACAGAAAGATATTCTAACTGTTTCAAGATGATTGTCAT 2531  
Db 2326 GGGTGATCTATGTTCTGTTTACCAGAAAGATATTCTAACGGGTTTAAAGATGCAATCTCAT 2385  
QY 2532 GGTGGCGGTCAATTACTGTCATACCTAAACGGGTGCAATTCAAAGGTTCTGCACTCTG 2591  
Db 2386 GGTGGAGATCTGTTTATGTGTACACCAAGTTAGCGGCTTCAAGGATCAGCTCCAATC 2445  
QY 2592 AATCTTCAGATCGTCTTCAACCAAGTGTGCTGGTGGGCTCTTGCTTAATGAGATCTTC 2651  
Db 2446 AATCTTCGATCGTCTCCATCAAGTTCCTGATGGGCGCTTGCGTGGTTGAGATTTTC 2505  
QY 2652 TTCAGCAATCATGGCCCTCTTGTGATGGGTATGGTGGCGGTCTGAATTTTGAAGA 2711  
Db 2506 TTGAGTAGGCAATGTCCTATTGTTGATGTTATGAGGTGGTGAATGGCTTGAGCGG 2565  
QY 2712 TTTTCTACATCAACTCCATCGTGTATCCTTGACATCTATTCCCTCTTGACTACTGT 2771  
Db 2566 TTGTCTACATTAACTCTGTGTGTTTACCCGTGACCTCTCTACCGCTCATCGTTTACTGT 2625  
QY 2772 ACATTGCCCTGCCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGCTGAATATGTT 2831  
Db 2626 TCTCTCCCTGCCATCTGTTCTTCTCACTGAAATTCATCGTTCCGAGATTAGCAACTAT 2685  
QY 2832 GCCAGCCTGTGTTCAATGTCATCTTTTATCTGCAATTTTGTCTACGAGCATCTAGAATG 2891  
Db 2686 GCGAGTATCCTCTTCAATGCGGCTCTTCTCGTCAATGCAATAACGGGTATCTCGAGATG 2745  
QY 2892 AGATGAGTGTGTGAATTGATGACTGTGAGGAATGAGCAGTTCTGGGTCATTGGA 2951  
Db 2746 CAATGGGGCAAAAGTTGGATCGATGATGTGGAGAAAACGACAGTTTGGGTCATTGGA 2805  
QY 2952 GGTGTGCTCTACACCTCTTGTGCTGTGTGCCAGGAACTTCTCAAGGTCATAGCTGTGTT 3011  
Db 2806 GGTGTTCTGCGCATCTGTTGCTCTCTTCCAAGGTCCTCTCAAGGTTCTTGCTGTGTC 2865  
QY 3012 GATACAGCTTCAACCGTGACATCAAAAGGGTGAATGATGAGAGTTCTCAGAGCTATAT 3071  
Db 2866 GACACTAACTTCAAGTCACATCAAAAGCAGCTGATGAGAGTTCCTGACCTTAC 2925  
QY 3072 ACATTCAATGACTACTTATTGATACCTCTTACCACTTGCTTCTAATGAACTTCATT 3131  
Db 2926 CTCTTCAATGACTTCACTTCTCATCCCTCCAAATGACTCTACTCATATAAGTTCATT 2985  
QY 3132 GGTGTGCTGCTGGCGTTTCAATGCGATCAATAACGATATGAGTCATGGGCCCCCTC 3191  
Db 2986 GGAGTCATAGTCGAGCTCTGATGCCATCAGCAATGAGATAGACTCGTGGGACCGCTT 3045  
QY 3192 TTTGGAACTATTCTTTGCAATTTGGGTGATGTGTCATTATCCCTTCTCAAAGGT 3251  
Db 3046 TTGGAAGACTGTTCTTTGCACTTTGGGTCAATTCATTACCCGTTCTTAAAGGT 3105  
QY 3252 TTGGTTGAAGGCAAAACAGGACACCAACGATTGTCATGCTGTGTCATTCTGCTGCT 3311  
Db 3106 TTGCTTGGAAACAAAGATAGAATGCAACCATTAATGTGCTGTGTCATCTCTGCGC 3165  
QY 3312 TCAATCTTCTGCTCTTGGGTTCGATGATCTTCTTGCGAAGGATGATGTCG 3371  
Db 3166 TCGATTCTTACACTTCTTGGGTCCGGTTAATCCGTTTGTGGCGA--AGGCGGTCT 3222  
QY 3372 CTCTTGAGAGAGTGTGTTGATTGC 3398  
Db 3223 ATTCTGAGATCTGTGTTAAGACTGC 3249

RESULT 10  
AAC83798

ID AAC83798 standard; cDNA; 3444 BP.  
XX  
AC AAC83798;  
XX  
DT 07-MAR-2001 (first entry)  
XX  
DE Arabidopsis thaliana cellulose synthase Cels cDNA.  
XX  
KW Arabidopsis thaliana; aspen; Cels; cellulose synthase; wood;  
KW mechanical stress response element; MSRE; cellulose biosynthesis; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN W020071670-A2.  
XX  
PD 30-NOV-2000.  
XX  
PF 18-MAY-2000; 2000WO-US13637.  
XX  
PR 21-MAY-1999; 99US-0135280.  
XX  
PA (UNMT ) UNIV MICHIGAN TECHNOLOGICAL.  
XX  
PI Chiang VL, Wu L, Chandrashekar JP, Carraway DT;  
XX  
DR MPI; 2001-032025/04.  
XX  
PT Novel polynucleotide encoding cellulose synthase useful for enhancing  
PT cellulose biosynthesis, modifying lignin biosynthesis, and for  
PT accelerating growth of plants and for upregulating cellulose synthase  
PT levels -  
XX  
PS Claim 8; Fig 7; 65pp; English.  
XX  
CC The present sequence is given in a specification relating to  
CC polynucleotides encoding cellulose synthase enzymes from Populus  
CC tremuloides and Arabidopsis. The cellulose synthases are designated  
CC PtcEla and AraxEla respectively. The polynucleotides may be introduced  
CC into plants where they are useful for altering the growth of a plant.  
CC Vectors comprising a cellulose synthase promoter, or a fragment of the  
CC promoter containing one or more mechanical stress response elements  
CC (MSRE), operably linked to a cellulose synthase gene are useful for  
CC causing stress-induced gene expression in a plant cell. A cDNA encoding a  
CC protein that binds to a positive MSRE may be delivered into the plant to  
CC increase expression of cellulose synthase. A cDNA in an antisense  
CC orientation may be delivered to reduce expression of cellulose synthase.  
CC The polynucleotides enable the development of transgenic tree varieties  
CC having increased cellulose content, decreased lignin content and  
CC therefore improved wood fibre characteristics. The cellulose synthase  
CC promoters are useful for synthesising greater quantities of  
CC high-crystalline cellulose in plant and preferably in trees, which  
CC permits an increase in proportion of cellulose in transgenic plants,  
CC greater strength of juvenile wood and fiber and acceleration of overall  
CC growth rate.  
CC Note: This nucleotide sequence is given in the specification as  
CC SEQ ID NO: 5. SEQ ID NO: 5 is referred to in claim 8 but it is described  
CC as a polypeptide sequence.  
XX  
SQ Sequence 3444 BP; 903 A; 664 C; 860 G; 1017 T; 0 other;

Query Match 37.9%; Score 1489.8; DB 22; Length 3444;  
Best Local Similarity 67.3%; Pred. No. 0;  
Matches 2189; Conservative 0; Mismatches 1032; Indels 30; Gaps 5;

QY 160 GGCTGTCGCGGCTCCACCAACGCAAGCTGTCGTCATCCGCGGAGCGGATC 219  
Db 9 GGTTAATGCGCGGTTCTCACAACAGGAATGAGTTGTCTCAATTAATGCCGATGAGAAAG 68  
QY 220 CCGGCGCGAAGCCGCGCGGAGAGAAAGGAGAGTGTGCGAGATTGCGGAGAGAG 279  
Db 69 CCCGAATAAGATCAAGTCCAAAGAGCTGAGTGACAGACATGTCAAATCTGAGAGATGAGA 128  
QY 280 TCGGCTTGCCCCCGGCGGAGACCCCTTGTGCGTGCACAGAGTGCGGCTTCCCGTCT 339

Db 129 TCGAATTGACTGTGATGGAGAACCGTTTGTGGCATGTAAACGAATGTGCATTCCCTGTGT 188  
Qy 340 GCCGGGACTGCTACGAATACGAGCCGGGAGGGCAGCAGAACTGCCCCAGTGCAAGA 399  
Db 189 GTAGACTTGTATGAGTACGAAAGACGAGAAAGCAATCAAGCTTGTCCACAGTGCAAAA 248  
Qy 400 CTGATACAAAGCGCCTCAAGGGGTGCCAACTGTGACCGGTGACGAGGAGAGAGACGGCG 459  
Db 249 CCGGTTTCAACGCTTAAAGGAAGTCCAAGAGTTGAAGGTGATGAAGAGAGAGATGACA 308  
Qy 460 TCGATGACCTGACAAACGAGTTCAACTGGGACGGCCATGACTCGAGTCTGTGCGCAGT 519  
Db 309 TTGATGATTTAGACAATGAGTTGA---GTATGAAATAATGGGATTGGATTGTGATCAGG 365  
Qy 520 CCAATGCTTACGGCCACATGAGCTACGGCCGTGGAGGTGACCCCTAATGGCGCCACAAG 579  
Db 366 TTTCTGAAGGTATGTCAATCTCTCGTCGCAACTCCGGTTTCCACAATCTGATTTGGATT 425  
Qy 580 CTTTCCAGCTCAACCCCAATGTTCCACTCTCACCAACGGGCAAAATGTTGATGACATCC 639  
Db 426 CAGCTCCACTGGCTCTCAGATTCCATTGTGACTTACGGCGACGAGGACGTTGAGATTT 485  
Qy 640 CACCGAGCAGCAGCGC---TGGTCCCTTCTTTCATGGGTGGTGGGAAAGAGATAC 696  
Db 486 CTTCTGATAGACATGCTCTTATGTCTCTCTTCACTTGGTGTCAATGCAATAGAGTTC 545  
Qy 697 ATCCCTTCTTATGCGGATCCAGCTTACTGTGCAACCCAGTCTATGACCACCA 756  
Db 546 ATCTGTCTTCTTCTTGACCCCGACCGTGGCTGCACATCGAAGCTGATGGTACCTCAGA 605  
Qy 757 AGGATCTTGTCTGCATATGGGTATGGTATGTGCTTGAAGAACGGATGAGAATTGGA 816  
Db 606 AAGATCTTGGGTTTATGTTATGGAAGTGTGCTTGAAGATCGGATGGAAGAAATGGA 665  
Qy 817 AGCAGAGACAAGAGAGATGCACAGACGGGGA-----ATGATGGTGGTGTATG 867  
Db 666 AGAGAAAGCAGAAATGAGAAACTTCAGGTGTAGGCATGAAGAGATCCTGATTTGAG 725  
Qy 868 ATGTGACGATGCTGATCTACCACTAATGATGAAGCAAGCAACAATGTCCAGAAAA 927  
Db 726 ATGTGATGATGCTGATTTCCAAATGATGATGAGGAAAGCAGCCATGTCTATGAAGA 785  
Qy 928 TTCCACTTCCATCAAGCCAGATTAATCCATATAGATGATTAATTCGGCTGTG 987  
Db 786 TACCAATCAATCGAGCAAGATAATCCTTACCAGATGTTAATGTGCTACGCTTGTGA 845  
Qy 988 TTTTGGGTTCTTCTTCCACTACCGAGTGCATCCGGTGAATGATGCATTGCTTGT 1047  
Db 846 TTTTGGTCTCTTCTTCACTACCGTATTTCTCACCCCGTCAAGATGCATATGCTTGT 905  
Qy 1048 GGTCTATCTGTATCTGTGAAATCTGTTTGGCAATCTTGGATTTCTGATCAATTCC 1107  
Db 906 GGCTTATTTCTGTATATGAGATATGTTGTGCTTTCATGGGTTCTTGATCAGTCC 965  
Qy 1108 CAAAGGTTCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTGACA 1167  
Db 966 CTAATGTAACCTATCGAGCGAAGACGTACTTGACCGACTCTCATTAAGATATGAGA 1025  
Qy 1168 AGAAGGCCAGCCATCTCAACTTGTCCAATTGATTTCTTGTCAAGTACGTTGATCCCT 1227  
Db 1026 AAGAAAGGAAAAACCGTCGGGACTATCCCTGTGATGATATTTGTTAGTACAGTGCAT 1085  
Qy 1228 TAAAGGAACCTCTTGTGTACAAACAATACTGTTCTATCTATCTTTCGGTGAATATC 1287  
Db 1086 TGAAGAGCCTCCGCTTATTACTGCAAAATACTGTGTTCTAATTTCTGTGCTGATATAC 1145  
Qy 1288 CTGTGTAAAGGTTTCTGTATGTTTCTGATGATGTGCTGCAATGCTAACGTTGAAG 1347  
Db 1146 CTGTGATAAAGGTGCTGTGTACGTAATCTGATGATGTGCTGCTAATCTTTGCAAG 1205  
Qy 1348 CATTAATGAACAATCTGAATTTGCAAGAAATGGGTTCTTCTGCAAAACGTAACAATA 1407

Db 1206 CTCTTTCTGAGACCGCTGAATTCCGAAGAAATGGGTTCCTTTCTGCAAGAAATATTGTA 1265  
Qy 1408 TTGAACCTCGGCTCCAGAGTGTACTTCCAAACAGAAATAGACTACTTGAAGAACAAG 1467  
Db 1266 TTGAGCCTCGTGTCCCGAATGGTATTTCTGCCATAAAATGACTACTTGAAGATAAAG 1325  
Qy 1468 TGGCAGCAACTTTGTTAGGGAGAGAGAGCAATGAAGAGAGATAGAGAAATTCAAAG 1527  
Db 1326 TTCAATCCGCATTTGTTAGGGAGCGCGAGCCATGAAGAGAGATTATGAAGATTCAAG 1385  
Qy 1528 TGGAATCAATGCTTAAGTTGCCAAAGCCCAAGAAAGTTCTGAAAGAGATGCAATGC 1587  
Db 1386 TAAAGATCAATGCTTTAGTAGCAACAGCACAGAAAGTGCTGAGGATGTTGACTATGC 1445  
Qy 1588 AAGATGAAACCCCTGGCCTGAAACAATGTTGATCATCTGGAATGATCAAGTCT 1647  
Db 1446 AAGACGGTACACCTTGGCCCGGTAATAGTGTGCGAGATCATCTGCGATGATCAGGCT 1505  
Qy 1648 TCCTTGGCCAAAGCGAGGCTTGACTGTGAGGAAATGAACCTGCCAGATTGCTTATG 1707  
Db 1506 TCCTTGAAGTGAACGGTGTCTGTGATGTGCAAAACAAGAGTTGCCCTCGATTAGTTACG 1565  
Qy 1708 TTTCTAGAGAGAAACGACAGGCTATTAACCATTAAGAAAGCTGTGCTATGAATGCAT 1767  
Db 1566 TTTCTGTGAGAAAGACCCGGATTGTATCACCATTAAGAAAGCTGAGCTATGAATTC 1625  
Qy 1768 TGGTCCGAGTCTCTGCTGTACTAACAATGCTCCATTTGTTAACTTGATGTGATC 1827  
Db 1626 TGATACGAGTCTTGGGTTCTATCAAAATGCTCTTACCTTGAATGTGATGTGATC 1685  
Qy 1828 ACTACATCAACAACAGCAAGGCTATTAAGGAAGCAATGTTTATGATGACCTTTAC 1887  
Db 1686 ACTACATCAACAATAGCAAAAGCTCTTAGAAGAGCAATGTGTTTCATGATGATCTCAGT 1745  
Qy 1888 TAGAAAGAAGGTTGCTATGTACAGTTCCTCAAGAAATTTGATGGATGATTCGCCATG 1947  
Db 1746 CAGAAAGAAATCTGTATATGTTCAGTTCCTCAAGGTTGCATGGGATGATGACACG 1805  
Qy 1948 ACCGATATGCTAACCGAATGTTGTCTTTTGTATATCAACATGAAAGGTTGGATGTA 2007  
Db 1806 ATCGATACTCAAAATCGCAATGTGTGTTCTTGTATATCAATATGAAGGTTGGATGGC 1865  
Qy 2008 TTCAGGCTCCAATTATATGTTGTACTGTGATGTGTAATTAAGAGCAGGCATTAATGCTT 2067  
Db 1866 TACAAGGCCATATATACGTGGTACAGTTGTGTTTCAAGAGCAGCGCTTTACGGAT 1925  
Qy 2068 ATGATGCCCCCAAAACAAGAACCAACATCAAGCACTTGCAACTGTGGCCCAAGTGT 2127  
Db 1926 TTGATGCAACGAAAGAAAGAGGGCCACGTAAGACATGCAATTGCTGGCCAAATGCT 1985  
Qy 2128 GCTTTTGTGCTGTGCTTGGCAATAGAGAGCAAAAGAAAGTACCAAAACCCAAACAG 2187  
Db 1986 GTCTCTATGTTTGTGTTCAAGAAAGATCGTAAAGCAAGACAGTGGCTGGATTAAGA 2045  
Qy 2188 AGAAGAAAAAGTTATATTTTCAAGAAAGAAAGAACCAATCCCTGCAATATGCTCTTG 2247  
Db 2046 AGAAGAAAGAA-----TAGGGAAGCGTCAAGCAAGATCCACGCAATTAAGAAAAATA 2093  
Qy 2248 GTGAATTGACGAAGCTGCTCCAGAGCTGAGAAAGAAAGCCGGTATTTGTAATCAAC 2307  
Db 2094 TCGAAGAGGGCCGCGTCAATAAGTTCTTAACGTAGAACAGTCAACCGAGGCAATGCAAA 2153  
Qy 2308 AAAAATTAGAAAAAATTTGGCCAATCTTCTGTTTGTTAATCCACACTTCTCGAGA 2367  
Db 2154 TGAAGTTGAGAGAAATATGGGCAAGTCTCCTGTATTTGTGATCTGCGCGTTCGAGA 2213  
Qy 2368 ATGTGGAACCTTGAAGTGAAGTCCCTGCTCTTTTGAAGAAGCTATACATGTCA 2427  
Db 2214 ATGCTGGGATGCTGAAGACCAAGCCCGCTTGTCTCTTAAGAAGCCATCCAAAGTCA 2273  
Qy 2428 TTAGTTGTGTTATGAAGACAAGACAGACTGGGGAAGAAAGATTTGGCTGATCTATGAT 2487  
Db 2274 TTAGTCGGGATATGAAGATTAAACTGAATGGGGAAGAAAGATTTGGGTGATCTATGCTT 2333

QY 2488 CAGTACAGAGATATTTCTAACTGGTTTCAAGATGCATTGTGATGTTGGCGGTCATTT 2547  
DB 2334 CTGTTACCGAAGATATTTCTTACGGGTTCTAAGATGCATTCTCATGGTTGAGACATGTTT 2393  
QY 2548 ACTGCATACCTAAACGGGTTGCATTCAAAAGGTTCTGCACCTCTGAATCTTTACAGTCGTC 2607  
DB 2394 ATTGTACACCAAGTTAGCGGCTTTCAAGAGATCAGCTCCAATCAATCTTTGCGATCGTC 2453  
QY 2608 TTCACCAGGTGCTCGGTGGGCTCTGGGCTTATTTGAGATCTTCTTCAAGATCATTTGCC 2667  
DB 2454 TCCATCAAGTCTTCGATGGGCGCTTGGTGGTGAATTTCTTGAAGGCAATGTC 2513  
QY 2668 CTCTTGGTATGGGATGGTGGCGGCTGAAATTTTGAAGATTTTCTTACATCACT 2727  
DB 2514 CTATTGGTATGGTTATGAGGTTGGGTTGAAATGCTTGAGCGGTTGCTTACATTA 2573  
QY 2728 CCATCGTATCTTGGACATCTATTCCTTGGCTTACTGATGATGCTGCTGCTGCT 2787  
DB 2574 CTGTGTTACCGGTGACCTCTCTACCGCTCATGCTTACTGTTCTCTCCCTGCT 2633  
QY 2788 GTTATTGACAGGAATTTATCACTCCAGAGCTGAATATGTTGCCAGCTGTGTTCA 2847  
DB 2634 GTCTTCTACGTGAATAATTCATCGTCCCGAGATGACAACTATGCGATATCCTCTTCA 2693  
QY 2848 TGTCACTTTTATCTGATTTTGTCTACGAGCATCTGAATGAGATGAGTGTGTG 2907  
DB 2694 TGGCGCTCTTCGTCGATTCGATTAACGGGTATCTCGAGATGCAATGGGCAAAAGTTG 2753  
QY 2908 GAATTGATGACTGTGAGGAATGAGCAGTTCTGGTCAATTGAGGTGTTCTCACACC 2967  
DB 2754 GGATGATGATTTGGTGAGAAACGACAGTTTGGTCAATTGAGGTGTTCTGCGCATC 2813  
QY 2968 TCTTGTCTGTGTTCCAGGACTTCTCAAGGTCAATAGCTGTGTGATACAAGCTTCACC 3027  
DB 2814 TGTGCTCTCTTCCAGGTTCTCTCAAGGTCTTGTGCTGTGCTGACACTTAATCTTCA 2873  
QY 3028 TGACATCAAAAGGTGAGATGATGAGAGTTCTCAGAGCTATATACATTCAAATGAGCTA 3087  
DB 2874 TCACATCAAAAGCAGTATGATGAGAGTTCTGACCTTTACCTTCAAAATGAGCTT 2933  
QY 3088 CTTTATGATACCTCTACCACTTGTCTTATGAATTCATTGTTGTTGCTGCTGCG 3147  
DB 2934 CACTTCTCATCCCTCAATGACTCTACTCATCATTAACGTCATTTGAGTCAATGCGAG 2993  
QY 3148 TTCAATGCGATCAATAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTATTCT 3207  
DB 2994 TCTGTATGCCATCAGCAATGATAGACTCGTGGGAGACCGCTTTCCGGAAGCTGTTCT 3053  
QY 3208 TTGCATTTGGGTGATGTTCCATCTTATCCCTTCTCAAAAGTTGGTTGGAAGGCAAA 3267  
DB 3054 TTGCATTTGGGTGATGTTCAATCTTATCCCGTCTTAAAGTTTGTGTTGGGAAACAAG 3113  
QY 3268 ACAGACACCAACGATGTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3327  
DB 3114 ATGAATGCCAACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3173  
QY 3328 TTTGGGTTGGATGATCTTCTTCTGCGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 3387  
DB 3174 TTTGGGTTGGGTTAATCGTTTGTGGCGA--AAGCGGCTCTATTTGAGATCTGTG 3230  
QY 3388 GTTTGATGTC 3398  
DB 3231 GTTTAGACTGC 3241

RESULT 11  
ABZ12754  
ID ABZ12754 standard; DNA; 3255 BP.  
XX AC ABZ12754;  
XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 559.  
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX OS Arabidopsis thaliana.  
XX PN WO200216655-A2.  
XX PD 28-FEB-2002.  
XX PF 24-AUG-2001; 2001WO-US26685.  
XX PR 24-AUG-2000; 2000US-227866P.  
XX PR 26-JAN-2001; 2001US-264647P.  
XX PR 22-JUN-2001; 2001US-300111P.  
XX PA (SCRI) SCRIPPS RES INST.  
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX PI Harper JF, Krepe J, Wang X, Zhu T;  
XX DR WPI; 2002-304127/34.  
XX PT Identifying a stress condition to which a plant cell has been exposed  
XX PT and producing plants with increased tolerance to these abiotic stresses  
XX PS Claim 144; SEQ ID NO 559; 577bp + Sequence Listing; English.  
XX CC The invention relates to identifying a stress condition to which a plant  
XX CC cell has been exposed, comprising:  
XX CC (a) contacting nucleic acid representative of expressed polynucleotides  
XX CC in the plant cell with an array or probes representative of the plant  
XX CC cell genome; and  
XX CC (b) detecting a profile of expressed polynucleotides in the plant cell  
XX CC characteristic of a stress response. The method is useful in the  
XX CC production of transgenic plants, cells and seeds and in producing plants  
XX CC with increased tolerance to abiotic stress. The present sequence is that  
XX CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
XX CC in methods of the invention.  
XX CC Note: The sequence data for this patent is not represented in the printed  
XX CC specification but is based on sequence information supplied to Derwent by  
XX CC the European Patent Office.  
XX SQ Sequence 3255 BP; 884 A; 621 C; 798 G; 952 T; 0 other;  
Query Match 36.4%; Score 1431.8; DB 24; Length 3255;  
Best Local Similarity 66.0%; Pred. No. 0;  
Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;  
QY 144 ATGAGGCGAGCGCGGCTGTGGCGGCTCCCAACCGCAAGCTGCTGCTATC 203  
DB 1 ATGAATATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60  
QY 204 CGCCGCGAGCGGATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 263  
DB 61 AACGCGATGAGAGTGCAGATTAAGATCAGTACAAAGACTGAGTGGCAACATGTCAA 120  
QY 264 ATTGCGCGAGCAGCTGCGGCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 323  
DB 121 ATCTGTGAGATGAATGAATTAAGTTAGCAGTGAAGCTTTGTGCTTGAACGAA 180  
QY 324 TGGCGCTTCCCGTCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 383  
DB 181 TGGCGATTCGCGGTTGTAGACCATGCTATGATGATGAACGTGAGAGAAATCAAGCT 240  
QY 384 TGCCCGCAGTCAAGACTGATCAAGCGGCTCAAGGCGTGCACAGTGTGACCGGTGAC 443  
DB 241 TGTCTCAGTGCAAACTCGATACAAAGGATTAAGGTAGTCCACGGGTGATGAGAT 300  
QY 444 GAGGAGGAGCAGCGGCTGATGACTGGAACAAGCTTCAACTGGGACGCGCATGCTCG 503



Db 301 GATGAAGAAGAAGACATTGATGATCTTGAGTATGAGTTTGATCATGGGATGACCCT 360

Qy 504 CAGTCTGGCCGAGTCCATGCTCTACGGCCACATGAGCTACGGCGGTGAGTGACCCT 563

Db 361 GAAACATGCCGCTGAAGCCGCACTCTCTCAAGCCTTAACACCGGTCGTGGTGGATTGAT 420

Qy 564 AATGGCGCGCACAAAGCTTTCAGCTCAACCCCAATGTTCCACTCTCAACCAACGGCAA 623

Db 421 TCAGCTCCACCTG-----GCTCTCAGATTCTCTCTTTGACTTATTGTGAT 465

Qy 624 ATGTTGATGACATCCCAACGGAGCAGCAGCGCTGCTGCTCTTTGATGGGTGGG 683

Db 466 GAAGATGCTGATATGTATTCTGATCGTCACTCTTATCGTCCCTCTTCAACGGGATAT 525

Qy 684 GGAAGAAGATACATCCCTCTTATGGGATCCAGCTTACCTGTGCAACCCAGTCT 743

Db 526 GGGAAATCGCTTATCTGCAACCGTTTACAGATTCTTCTGCACCTTCAACGGCAGATCA 585

Qy 744 ATGACCCATCCAAGGATCTTGCTGCATATGGGTATGTTAGTGTGCTGGAAGAACGG 803

Db 586 ATGTTCTCAGAAAGATATTGCGGAATATGTTATGGAAGTGTGCTTGGAAAGACCGT 645

Qy 804 ATGAGAAATTGGAAGCAGACAGACAGAGAGATGCACCAAGCGGGAATGATGTTGCT 863

Db 646 ATGGAAGTTGGAAGAGACGACAGACGCGAAAGCTTCAAGTCATTAAGCATGAAGAGGA 705

Qy 864 GATGATGCT-----GACGATGCTGATCTAACCACTAATG 896

Db 706 AACAAATGTCAGGTTCCAAATGATGACGACGAACCTAGATGATCTTGCATGCTATGATG 765

Qy 897 GATGAAGCAGACAACAACCTGTCCAGGAAATTCACCTTCATCAAGCAGATTAAATCCA 956

Db 766 GATGAAGGAAGCAAACTCTCTCAAGAAAGCTAATTCGTTCAAGCAGAATAAATCCT 825

Qy 957 TATAGATGATATCATTTATTCGCTGTGGTTTGGGTTCTTCTTCCACTACCGAGTG 1016

Db 826 TACAGAGATGTTAATCTGTGTGCGCTCGCGAATCTTGCTTTCTTCATTATAGAATT 885

Qy 1017 ATGCATCCGCTGAATGATGCATTTGCTTGTGCTCATATCTGTATCTGTGAATCTGG 1076

Db 886 CTCCATCCAGTCAATGATGCATATGATGATTATGTTAAGCTCAAGTTATGCGAGATATGG 945

Qy 1077 TTTGCCATGCTTGGATTCTTGATCAATTTCCAAAGTGTTCCCTATTGAGAAGAGACT 1136

Db 946 TTTGCAGTGTCTTGAATCTTGATCAATTTCCCAATGTTATCTTATAGAACGTGAACA 1005

Qy 1137 TACCTAGACCGCTGTCACTGAGTTTCGACAGGAAGGCCAGCCATCTCAACTGTCTCA 1196

Db 1006 TACCTCGATGACTCTCTCTCAGTACGAGGAAGAAAGAAACCGTCAGAGATTAGCACCT 1065

Qy 1197 ATTGATTTCTTGTCACTACGTTGATCCCTTAAAGAACCTCCCTTGTGCAACAACAT 1256

Db 1066 GTTGATGTTTGTAGTACAGTGATCCGTTGAAGAGCCACCCCTGATTACAGCAAAAC 1125

Qy 1257 ACTGTTCTATCTCTTTCGGTGGATTATCTGTGATAGGTTTCTTGCTATGTTCT 1316

Db 1126 ACAGTCTTTCATTTCTAGCAGTTGATTATCTGTGATTAAGGTTGCGTGTATGTATCA 1185

Qy 1317 GATGATGCTGCAATGCTAACGTTTGAAGCATTTATGAAACATCTGAATTGGCAAG 1376

Db 1186 GACGATGTCAGCTATGCTTACATTGAACTCTCTCTGATACAGCTGAGTTTGTCTAGA 1245

Qy 1377 AAATGGTCTTCTTCTGCAAAACGGTACAAATATGAACCTCGCGCTCAGAGTGTACTTC 1436

Db 1246 AAATGGTCTTCTTCTTGAAGAGTTTAATATCGAACCCAGAGCTCTGAGTGTATTTT 1305

Qy 1437 CAACAGAAAGTACTACTTGAAGAACAAGTGGCAGCAAACTTGTGTAAGGAGAGAGA 1496

Db 1306 TCTCAGAAAGATGATTACTGAAACAACAAGTTCACTCTGTTTGTGAGGAAACGTCT 1365

Qy 1497 GCAATGAAGAGAGTATGAGGAATTCAAGTGAAGATCAATGCTTATGTTGCCAAAGCC 1556

Db 1366 GCTATGAAGAGAGATTATGAGGAGTTTAAAGTGAAGATAAATGCACGTGTTGCTACTGCA 1425

Qy 1557 CAGAAAGTCTCTGAAGAGATGCACAATGCAGATGGAACCCCTGGCTGGAACAAT 1616

Db 1426 CAGAAAGTCCCTGAGGAAGGTTGACTATGCAGAGATGGAACCTTGGCTTGAACAAC 1485

Qy 1617 GTTCGTGATCATCTTGAATGATTCAAGTCTTCTTGGCCAAAGCGGAGGCTTGACTGT 1676

Db 1486 GTCCGTGACCATCTGGAATGATTCAAGTGTCTTGGTTCATAGTGAGATTCTGTATACG 1545

Qy 1677 GAGGAAATGAAGTCCACAGATTGTTTATGTTTCTAGAGAAACGACAGCCTATTAAC 1736

Db 1546 GATGTTAATGAGTTACACAGTCTAGTATGTTTCTCGTGAAGACGGCTGATTTGAT 1605

Qy 1737 CATCATAAAGAAAGCTGTGCTAATGATGATGTTGTCGAGTCTCTGCTGACTAACAAAT 1796

Db 1606 CACCACAAGAAAGCTGAGCTATGAATCTTGTATCCGAGTCTCTGCTGTCTATCAAAAC 1665

Qy 1797 GCTCATATTTGTTAACTTGATGTGATCACTACATCAACAACAGCAAGCTATTAAG 1856

Db 1666 GCTCCTTAACCTTCTTAATGTGATGTGATCACTACATCAACAACAGCAATTTAGA 1725

Qy 1857 GAAGCAATGTGTTTATGATGACCCCTTACTAGGAAGAAAGGTTGCTATGTACAGTTC 1916

Db 1726 GAATCTATGTGTTTCAATGATGACCCGCAATCGGGAAGAAAGTTGTATGTTCAGTTT 1785

Qy 1917 CCTCAAGATTTGATGGATTGATCGCCATGACCGATATGCTAACCGGAATGTGCTTT 1976

Db 1786 CCGCAGAGATTTGATGGATGATGACATGATAGATACTCAAAACCGTAACGTTGTGTTTC 1845

Qy 1977 TTTGATATCAACATGAAGGTTTGATGTTATTCAGGTTCCAAATTTATGTTGTTACTGGA 2036

Db 1846 TTTGATATTAACATGAAGGTTCTGATGGGATACAAAGACCGATATATGTCCGGACAGGT 1905

Qy 2037 TGTGTTATTAAGGACGCAATTATGTTATGATGCCCCCAAAACAAAGAACCCACCA 2096

Db 1906 TGTGTTTGAAGACAGGCTCTTTATGCTTTGATGACCAAGAAAGAAACCAACCA 1965

Qy 2097 TCAAGACTTGCAACTGCTGCGCCCAAGTGTGCTTTGCTGTGCTGCTTGGCAATAGG 2156

Db 1966 GGCAAAACCTGTAACGTGTGCTTAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2025

Qy 2157 AAGCAAAAGAAAGTACCAAAACCAAAACAGAGAAAGAAAGTTATTTTCAAGAAA 2216

Db 2026 AGTAAACGAAAGCCAAAGATGAAGAAACCTAACACTAAAGAGACTTCA----- 2073

Qy 2217 GAAGAAACCAATCCCTGCAATATGCTCTTGTTGAATTTGACGAAGCTGCTCCAGAGCT 2276

Db 2074 -----AAGCAGATTTCATGCGCTAGAGATGTGCGACGAAGGTGTATCGTCCAGTGTCA 2127

Qy 2277 GAGATGAAGAAAGCCGCTATTTGAATCAACAATAATTGAAGAAATTTGGCCAATCT 2336

Db 2128 AATGTTGAAGAGAGATCTGAAGCAACAATTTGAATTTGAAGAAAGTTTGAACAATCT 2187

Qy 2337 TCTGTTTTTGTACATCCACTTCTCGAATGTGGAACCTTGAAGAGTCAAGTCTCT 2396

Db 2188 CCGGTTTTGTTGCTGCTGTGTTCTACAGAACGTTGAGTTCCCGTAAACGCAAGCCCC 2247

Qy 2397 GCTTCTCTTTGAAGAAAGCTATACATGTCTATTAGTTGTGTTATGAAGACAAAGACAGAC 2456

Db 2248 GCATGTTTGTTAAGAGAAAGCAATTCAAGTTATTAGCTGCGGTTACGAAGATAAAACGAA 2307

Qy 2457 TGGGAAAGAGATTGCTGATCTATGATCAGTTACAGAAAGATATTCTAATGTTTC 2516

Db 2308 TGGGAAAGAGATCGGCTGATTTATGATCGGTGACTGAAGATATCTGACGGGTTTC 2367

Qy 2517 AAGATGATTTGTCATGTTGGCGTCAATTTACTGCATACCTTAAACGGGTTGCATTCAA 2576

Db 2368 AAGATGATTTGCCATGATGAGATCTGTCTATGCTTAAAGCGTGAAGCTTTTAA 2427

Qy 2577 GGTTCGACCTCTGAATCTTACAGATGTTCTTCAACGAGTCTTGGTGGGCTCTTGGG 2636

Db 2428 GGATCTGCTCTATTAACCTTGCAAGTGTCTTCAATCAAGTTCTACGTTGGCTCTTGGC 2487

QY	2637	TCTATTGAGATCTTCTTCAGCAATCATTTGCCCTCTTTGGATATGGGATATGGTGGCGGCTCTG	2696
Db	2488	TCTGTAGAGATTCTTTGAGCAGACATTGTCCGATATGGTATGGTTATGGTGGCTTTTA	2547
QY	2697	AAATTTTGGAAAGATTTTCTTACATCAACTCCATCGTGTATCCTTGACACTTAATCCC	2756
Db	2548	AAATGTTGAGAGATTCTTACATCAACTCTGTGCTATCCTTGACTTCACCTCCA	2607
QY	2757	CTCTTGCGTTACTGTACATGCTCGCATCTGTTTATGACAGGAAATTTATCACTCCA	2816
Db	2608	TTGATCGTCTATGTCTCTCCCGCGGTTTGTACTCACAGAAATTCATCGTCCCT	2667
QY	2817	GAGCTGAATATGTTGCCAGCCTGTGTTTCATGTCACTTTTATCTGCATTTTTCGTAAG	2876
Db	2668	GAGATTAAGCACTACGCGAGTATACCTTTCATGTCTCATGTTCATATCCATAGCAGTAAC	2727
QY	2877	AGCATCCTAGAAATGAGATGAGTGTGTTGGAATTTGATGACTGTGTGAGGAATGACAG	2936
Db	2728	GGAATCCTCGAAATGCAATGGGGAGGTGTGGAATCGATGATTGTTGGAGAAACGACAG	2787
QY	2937	TTCTGGGTCATTGGAGGTGTCTCACACCTCTTGTCTGTGTTCAGGAGACTTCTCAAG	2996
Db	2788	TTTGGGTAATCGAGGGGCTCTCGCATCTATTGTCTGTCTTCAAGTTTGTCTCAAA	2847
QY	2997	GTCATAGCTGTGTTGATACAAGCTTCAACCGTGACATCAAGGGTGGAGATGATGAGAG	3056
Db	2848	GTTCTAGCCGAGTTAACGCAATTTACAGTCACTTCAAAAGCAGACAGATGAGCT	2907
QY	3057	TTCTCAGAGCTATATACATCAATGAGACTACCTTATTTGATACCTCCTACCACCTTGCT	3116
Db	2908	TTCTCTGAGCTTTACATCTTCAAGTGAGACAACCTTGTGATTCCTCCGACAACACTTCTG	2967
QY	3117	CTATTGAACCTCATTTGGTGTGCTCGCTGGCGTTTCAATGGCATCAATTAACGGATATGAG	3176
Db	2968	ATCATTAACATCATTTGAGTTATTTGCGGCTTCTGATGCCATTAGCAATGGCTATGAC	3027
QY	3177	TCATGGGGCCCCCTTTGGGAAGCTATTCTTTGCAATTTGGGTGATTTGCCATCTTTAT	3236
Db	3028	TCATGGGGACCTCTTTGGGAGACTTTTCTTGGCTCTTTGGGTCAATTGTTCAATTATAC	3087
QY	3237	CCCTTTCTCAAAAGTTTGGTTGGAAGCAAAACAGACACCAACGATTGTCAATCGTCTGG	3296
Db	3088	CCATTCCTCAAGGGAATGCTTTGGGAAGCAACAAATGCTTACGATTATTGTGTCTGG	3147
QY	3297	TCCATTCTGTGCTTCAATCTTCTCGCTCCTTTGGGTTGGATTGATCCTTTCCTTGGG	3356
Db	3148	TCTATTCTTCTAGCTTCGATCTTGACACACTCTGTGGGTCAAGTTAACCCGTTGTGGCT	3207
QY	3357	AAGGATGATGTCGCTTCTTTGAGAGAGTGTGTTGGATTG	3397
Db	3208	A--AAGGGGACCAGTGTGGAGATCTGTGGTCTGAATTG	3245

	RESULT 12
AAV06566	ID AAV06566 standard; cDNA; 3828 BP.
XX	
AC	AAV06566;
XX	
DT	06-JUL-1998 (first entry)
XX	
DE	Arabidopsis cellulose biosynthetic gene clone Ath-A.
XX	
KM	Cellulose; cellulose synthase; RSW1 gene; beta-1,4-glucan;
KW	transgenic plant; ss.
XX	
OS	Arabidopsis thaliana var. Columbia.
XX	
FH	Key Location/Qualifiers
FT	CDS 239..3490
FT	/tag= a
XX	
PN	WO9800549-A1.

XX 08-JAN-1998.  
PD  
XX  
XX 24-JUN-1997; 97WO-AU00402.  
PF  
XX  
XX 27-JUN-1996; 96AU-0000699.  
PR  
XX  
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (AUSU ) UNIV AUSTRALIAN NAT.  
XX  
XX Arioli A, Betzner AS, Peng L, Williamson RE;  
PI  
XX  
XX WPI; 1998-086974/08.  
DR  
DR P-PSDB; AAW33818.  
XX  
XX DNA encoding cellulose biosynthetic enzyme - useful for manipulation  
PT of cellulose and beta-1,4-glucan  
PT  
XX  
PS Claim 12; Page 123-130; 207pp; English.

CC cDNA clone Ath-A was isolated from an *Arabidopsis thaliana* cDNA  
CC library using probes produced by PCR with primers (see AAT93632-34)  
CC based on cellulose synthase RSW1 genomic clone 23H12 (see AAV06563)  
CC and EST clone AAT20782 (see AAV06562). It is closely related to RSW  
CC cDNA (see AAV06565) and to partial genomic clone 12C4 (see AAV06564)  
CC and encodes a 1084-polypeptide (see AAW33818). Claimed nucleic acid  
CC molecules (see AAV06562-69) coding for claimed polypeptides (see  
CC AAW33816-20 and AAW46202) involved in cellulose biosynthesis can be  
CC used to manipulate the cellulose and/or beta-glucan content of  
CC transgenic plants. Expression of nucleic acids in the sense  
CC orientation increases the level of cellulose and reduces the level  
CC of non-crystalline beta-1,4-glucan and starch, providing plants  
CC with modified strength and/or shape and/or fibre properties, or  
CC having increased resistance to stresses or pests. Antisense,  
CC ribozyme or co-suppression molecules can be used to reduce the  
CC cellulose content of a transgenic plant, e.g. to improve  
CC digestibility or to alter carbon partitioning such that increased  
CC carbon is available for growth, rather than deposited as cellulose.

Sequence 3820 BP; 1056 A; 721 C; 912 G; 1139 T; 0 other;

Query Match	36.3%	Score 1429.4;	DB 19;	Length 3828;
Best Local Similarity	65.8%;	Pred. No. 0;		
Matches 2172; Conservative	0;	Mismatches 1066;	Indels 63;	Gaps 4;

[illegible]





```

Db 2766 ATGTTATGGTGGTGGTTAAATAGTTGAGAGATTCTTACATCAACTCTGTCCT 2825
OY 2737 ATCCTTGACATCTATTCCTCTTGCTTACTGATCATTCGCTGCCATCTGTTATTGA 2796
    |||||
Db 2826 ATCCTTGACATCTTCCATTGATCGCTATTGTTCTCTCCCGGTTGTTACTCA 2885
OY 2797 CAGGAAATTATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGTTCACTGTCATT 2856
    |||||
Db 2886 CAGAAATTATCTCTCCAGATTAAGCACTACGAGTATCTCTTCACTGTCATGT 2945
OY 2857 TTATCTGCATTTTCTCTACGAGCATCTAGAAATGAGATGAGTGGTGGTAATTGATG 2916
    |||||
Db 2946 TCATATCCATAGCACTAAGTCCGAAATGCAATGGGAGGTGCGAAATCGATG 3005
OY 2917 ACTGGTGAGAAATGACAGTCTGAGTCAATTGAGGTGTCTCCTACACCTCTTGCTG 2976
    |||||
Db 3006 ATTGGTGAGAAACGACAGTTTGGTAATCGAGGGGCTCTCGCATCTATTGCTC 3065
OY 2977 TGTTCAGGACTTCTCAAGTTCATAGCTGTGTGATACAGCTTCACCGTGACATCAA 3036
    |||||
Db 3066 TGTTCAGGTTTGTCTCAAAGTTCTAGCCGAGTTAAACGAAATTCACAGTCACTCAA 3125
OY 3037 AGGTGAGATGATGAGAGAGTCTCAGAGCTATATACATTCAATGAGTACTTATTGA 3096
    |||||
Db 3126 AAGCAGCAGACGATGAGAGCTTCTGAGCTTTACATCTCAAGTGAGCAACTTGTGA 3185
OY 3097 TACCTCCTACACCTTGCTTCTATTGAATTCATGTGTGTGCTGCGCTTCAAAATG 3156
    |||||
Db 3186 TTCCTCCGACAACTTCTGATCATTAACATCATGAGTTATTGTGCGCTTCTGATG 3245
OY 3157 CGATCAATAACGATATGATGATGAGGCCCCCTTTGGGAAGCTATCTTGCAATTT 3216
    |||||
Db 3246 CCATTAGCAATGAGCTATGACTGAGGACCTCTTTGGAGACTTTCTGCTCTTT 3305
OY 3217 GGGTGATGTCATCTTATCCCTTCTCAAGGTTTGGTTGGAAGCAAAACAGACAC 3276
    |||||
Db 3306 GGGTCATGTTTATTAATACCATTCCTCAAGGAAATGCTTGGAAAGCAAAATGC 3365
OY 3277 CAACGATGTCATCGTCTGCTCATTCTGCTGCTTCAATCTTCTGCTCTTGCGTTC 3336
    |||||
Db 3366 CTACGATTATGTGTGTCTGCTATTTCTTCTAGCTTGAATCTTGTGGGTCA 3425
OY 3337 GGAATGATCCTTCTCTGCGAAGGATGATGCTCCGCTTCTTGAGAGTGTGTTGATT 3396
    |||||
Db 3426 GAATTAACCCGTTGTGCTA--AAGGGGAGCCAGTGTGAGATCTGTGCTGAATT 3482
OY 3397 G 3397
    |
Db 3483 G 3483

```

RESULT 13  
 AAA67114  
 ID AAA67114 standard; DNA; 3851 BP.  
 XX  
 AC AAA67114;

XX 31-OCT-2000 (first entry)  
 DE Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.  
 XX  
 KM Eucalyptus grandis; pinus radiata; Monterey pine; modification;  
 KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;  
 KM transgenic plant; ds.  
 XX  
 OS Pinus radiata.  
 XX  
 PN WO200022092-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 08-OCT-1999; 99WO-NZ00169.  
 XX

```

PR 13-OCT-1998; 98US-0170862.  

PR 11-AUG-1999; 99US-0148426.  

XX  

PA (GENE-) GENESIS RES & DEV CORP LTD.  

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  

XX  

PI Bloksberg LN;  

XX  

DR WPI; 2000-339328/29.  

DR P-PSDB; AAB16307.  

XX  

PT New genes encoding proteins involved in a plant polysaccharide  

PT biosynthetic pathway, useful for modulating or altering the  

PT polysaccharide content, composition or structure of the plant  

XX  

PS Claim 1; Page 71-72; 301pp; English.  

XX  

CC The present invention describes isolated polynucleotides (PN) comprising  

CC a sequence selected from one of 835 nucleotide sequences given in  

CC AAA67073 to AAA67907, their (reverse) complements, sequences producing  

CC an Expectation (E) value of 0.01 or less compared to the 835 sequences,  

CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the  

CC 835 sequences or sequences that are degenerately equivalent or allelic  

CC to the 835 sequences. The polynucleotides are used to modify the  

CC activity of a polypeptide involved in a polysaccharide biosynthetic  

CC pathway in the plant. They are especially used to modulate or alter the  

CC polysaccharide content, composition or structure of the plant. AAB16268  

CC to AAB16340 are proteins encoded by some of the polynucleotide sequence  

XX given in the present invention.  

SQ Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 other;  

XX  

Query Match 35.9%; Score 1411.4; DB 21; Length 3851;  

Best Local Similarity 66.3%; Pred. No. 0;  

Matches 2206; Conservative 0; Mismatches 1016; Indels 105; Gaps 8;  

OY 123 GGAAGTGAGGGGAGGAAGCGATGAGCGCGCGGCTGTGCGCGCTCCCAAC 182  

    |||||  

Db 132 GGCAGAGAAATGAATTAACAATGAGAACCGCGGCTTGTGCGGTTCTATAAC 191  

OY 183 CGCAACGAGCTGTCATCCGCGCGAGCGGATCCCGGCGGAAGCCGCGGGAG 242  

    |||||  

Db 192 AGAACGAGTTCGTGTCATCC-----ATGACATGAGGAGCCGAAGCTTGAACAG 245  

OY 243 CAGAACGGGCGAGGTGTGCCAGATTGCGGCGAGCAGCGTCCGCTTGGCCCCGGGGAG 302  

    |||||  

Db 246 TTGAGTGCCACGCTCTGCCAGATTGTGGCGAGGAGCGTGGGCTTAAACAGACGGCGAG 305  

OY 303 CCTTCGTGCGGTGCAACGAGTGCCTTCCCGCTGTGCGCGGACTGCAATACGAG 362  

    |||||  

Db 306 CTGTCGTTCTGTATAGAGTGGGTTTCTGTGTGCGCGGTGCTATGATACGAG 365  

OY 363 CGCCGGAGGCGACGAGAACTGCCCCAGTGCAAGACTGATCAAGCGCTCAAGGC 422  

    |||||  

Db 366 AGACGAGAGGAATACAGTCGTGCCCGCAGTGCAATACCTGTTACAAGCGTCAAAAAGGG 425  

OY 423 TGCCAACGTGTGACCGGTGACGAGAGGAGGAGCGGCTGATGACTGGAACAAGATTG 482  

    |||||  

Db 426 AGTCCACGGGTGGAAGGTGACGATGATGAAGAAGCGTTGATGACATGAATTT 485  

OY 483 AACTGGAGCGCCA-----TGACTCGAGTCTGTGGCGGAGTCACTCTACGGCAC 536  

    |||||  

Db 486 AATGTGAGACTCAACAAAGAAACAGGACGACATCACCGAGCGATGCTCCACGAGCG 545  

OY 537 ATGAGCTACGGCGGTGAGGTGACCTTAATGGCGCG-----CACAACTTTCAGCTC 590  

    |||||  

Db 546 ATGAGCTATGGCCGAGGTCCCGACGACGAAATTCGAGATTGCTCATTAATCCAGAGCTT 605  

OY 591 AACCCCAATGTTCACTCTCAACCAAGGCA--AATGTGATGACATCCACCGGAG 647  

    |||||  

Db 606 CCTCCGAGATTCTCTACTTGCAACGCGCACTCGTTGTGAGTGGGAGATTCCACG 665  

OY 648 CAGCAGCGCTGTGCTTCTTTCATGGGTGTGGGGAAGAGATACATCCCTTCT 707

```

Db 666 TCATACTACGACAGACACCAATTGCTTGCCAACTCTGCAATGCTGAAGCGTGTGCATCCA 725  
Qy 708 TATGCGGATCCCAAGCTTACCTGNGACAACCCAGGTCTATGACCCTCAAGATCTTGCT 767  
Db 726 AGCTCCGAGCCGGGG-----AGTGAAGGATCATGTGATCCAAACAGGATATTGGT 779  
Qy 768 GCATATGGGTATGTAGTGTGCTTGGAAGGAACGGATGGAGAAATTGGAAGCAGAGACAA 827  
Db 780 TCTTATGGCTTTGGAACTGTCTTGGAAAGGAGCGAGCGATGGTTATAATCGAAGAA 839  
Qy 828 GAGAGATGCACCGACGCGGAATGATGSGTGG----- 862  
Db 840 AACAAATCAGGCCAGTTGATATGACGGAAGGAGATATCAATATATGCGGGGTTGCA 899  
Qy 863 -----TGATGATGATGATGATGCTGATCTACCACTAATGATGAAGCAAGACAA 911  
Db 900 CCAATGAGCCTGAAGATTATATGATCCGATATGCCAATGACCGATGAAGCAAGCAG 959  
Qy 912 CAACGTCCAGGAAATTCACCTTCATCAAGCCAGATTATCCATATAGATGATATC 971  
Db 960 CCACTGTCCCGAAAGTGCCTTCTCAAGCAAAATAATCCATACCGAATGCTATT 1019  
Qy 972 ATATTCGGCTTGTGTTTGGGTTCTTCTTCCACTACCGAGTGAATGTCATCCGGTGAAT 1031  
Db 1020 GTAAATTCGACTGATAGTGTGGTATTTTCTCCGCTATCGTCTCTGAATCCAGTGAAG 1079  
Qy 1032 GATGCAATTTGCTTTGTGGCTCATATCTGTTATCTGTGAATCTGTTGCCATGCTTGG 1091  
Db 1080 AATGCATATGGGCTCTGGGCCACTTCTATCGTTGTGAATCTGTTGCCCTGTGCAATGG 1139  
Qy 1092 ATTCGTGATCAATTCGCAAGTGTCTCCCTATTGAGAGAGAGACTTAACCTAGACCGGCTG 1151  
Db 1140 ATCTTGATCAGTTTCCCAAGTGTGCTTATCAGTCGTGAACGTATCCTGATCGACTG 1199  
Qy 1152 TCACTGAGGTTGACAGAAGGCGCCACCTCACTCACTTGTCCAATTGATTTCTTGTC 1211  
Db 1200 TCATTAAGGTACGAACGAGAAGGCAACCATCAATGCTTGCACTGTTGACTCTTGTG 1259  
Qy 1212 AGTAGGTTGATCCCTTAAGGAACCTCTTGTGTCACAACAATACTGTTCTATCTATC 1271  
Db 1260 AGTACTGTAGATCCACTGAAGAGCTCTCTTGGTTACTGCAATACAGTATTATCAATC 1319  
Qy 1272 CTTTCGGTGAATTATCCTGTGTGATGAAGTTTCTGTATGTTTCTGATGATGCTGCA 1331  
Db 1320 CTTTCAGTAGACTACCTGTAGACAATGTCTCTGTATGTCCTGATGACGAGCGTGG 1379  
Qy 1332 ATGCTAACGTTTGAAGCATTTGAACATCTGAATTTGCAAGAAATGGTTCCTTTC 1391  
Db 1380 ATGCTTACTTTTGAATCTCTCTGAGACCTCAGAATTTGCCAGAAATGGTACCAATTC 1439  
Qy 1392 TGCAACCGGTACAATATTGAACCTCGCGCTCCAGAGTGTACTTCCAACAGAGATGAC 1451  
Db 1440 TGCAAGAAATTCGACATTTGAGCTTCGCGCTCCCGAAATCTATTTCTCTCAGAAATTGAC 1499  
Qy 1452 TACTTGAAGACAAAGGTGGCAGCAAACTTTGTTAGGAGAGAGACAATGAAGAGAGAG 1511  
Db 1500 TATCTGAAGGACAAATTTCAACCCACTTTGTCAAGAGCGCGCTGCATGAAGAGAGAA 1559  
Qy 1512 TATGAGGAATTCAGGTTGAGATCAATGCTTAGTTGCCAAAGCCGAGAAAGTTCTGAA 1571  
Db 1560 TATGAAGAAATTCAGGTGCGCATCAATCGTGTGTTGCAAGGCTCTTAAAGTGCCCAAG 1619  
Qy 1572 GAAAGATGACAATGCAAGATGAACCCCTGCGCTGGAACAATGTTGATGATCAATCT 1631  
Db 1620 GAAAGATGACAATGCAAGAGCGTACGCTTGCGCTGTATATAATACCGTGACCAATCT 1679  
Qy 1632 GGAATGATTCAGGCTCTTGGCCAAGCGAGCGCTTGACTGTGAGGGAATGAATGACTG 1691  
Db 1680 GGTATGATCCAGTGTCTTGGGTACAGTGGCGCTTCGATACAGAAGGCAATGAGCTT 1739  
Qy 1692 CCACGATGCTTATGTTTCTAGAGAGAAACGACCAAGCTTAAACCATCATTAAGAAAGCT 1751

Db 1740 CCTCGGCTAGTATATGTTTCTCGTGAGAAAGACCTGGTTTCCAGCATCACAAAGGCC 1799  
Qy 1752 GGTGCTATGAATGCATTTGGTCCGAGTCTGTGCTGTACTAACAAATGCTCATATTGTTA 1811  
Db 1800 GGTGCTATGAATGCTTTGGTTCGGGTTTCTGCTGTCTCACCAATGCTCATTTATGCTG 1859  
Qy 1812 AACTTGATGTGATCACTACATCAACAACAGGAGGCTATTAAGGAAGCAATGCTTTT 1871  
Db 1860 AATCTGATGTGATCACTACATTAACAATGACAGGCAATCAGGGAAGCATGTGCTTT 1919  
Qy 1872 ATGATGACCCCTTACTAGAAAGAGTTTGCTATGTACAGTTCCTCAAGATTTGAT 1931  
Db 1920 ATGATGATCCTCAGGTTGGGAGAAAGTCTGTATGTCCAATTCCTCAGAGATTCGAT 1979  
Qy 1932 GGGATTGATCGCCATGACCGATATGCTAACCGGAATGTTGCTTTTGTGATATCAACATG 1991  
Db 1980 GGTATTGATCGCAATGACCGGTTACGCCAATCGAAACACCGTATTCTTTGATATCAACATG 2039  
Qy 1992 AAAGTTTGATGTTATTCAAGGTCCTCAATTTATGTTGCTACTGATGTATTTAGAAAG 2051  
Db 2040 AAAGTCTGATGAAATTCAGGGCTGTATATGTGGAACTGATGATGATGTTCAGAAAG 2099  
Qy 2052 CAGGCATTATATGTTATGATGCCCCCAAAACAAGAACCAACCATCAAGACTTGAAC 2111  
Db 2100 CAAGCTCTATATGGTATGCGCTCCCAAGGCCCAAAAC----- 2139  
Qy 2112 TGCTGGCCCAAGTGTCTTTTGTCTGTGCTGCTTGGCAATAGGAAGCAAAAGAACT 2171  
Db 2140 -----GTCCCAAGATGTGACCTGTGATGTCTCCCTGTGTGCGGTCTCTGTAAGAAGTCT 2195  
Qy 2172 ACCAAACCCAAAACAGAGAAGAAAAGTTATTTTCAAGAAAGAGAGAACCAATCC 2231  
Db 2196 CCGAAGAAAATATAGTACCAAGAAAAG-----TGACGAAATCCCAAGCT 2237  
Qy 2232 CCTGCATATGCTCTTGGTGAATTGACGAAGCTGCTCCAGAGCTGAGAAAGAAAGGCC 2291  
Db 2238 CCGGCTTACATCTGACGAGCGGATCGAGGAAGGATGAAGGTTATGATGACGAAAGAGCA 2297  
Qy 2292 GGTATTGTAATCACAAAAATTAGAAAAGAAATTGGCCATCTCTGTTTGTGTACA 2351  
Db 2298 TTGTTGATGAGCCAACTAGACTTCGAGAAGAGTTTGGCCAGTCTTCACTTTTGTCAA 2357  
Qy 2352 TCCACACTTCTCGAATAGTGTGAACCTTGAAGTGAAGTCCGCTCTCTTTTGAAA 2411  
Db 2358 TCCACTGTGATGAGAGATGTTGTTCCGCAACAGCAATCCAGCTGAATGTTGTAAG 2417  
Qy 2412 GAACTATACATGTCTATTAAGTTGTGTTATGAAACAAGACAGACTGGGGAAAAAGATT 2471  
Db 2418 GAGGCTATTATGATGATCAAGCTGTGATATGAAACAAGGATGGGAAAAAGAGCTT 2477  
Qy 2472 GGCTGATCTATGATCAGTTACAGAGATATTTCTAACTGTTTCAAGATGATGTCAT 2531  
Db 2478 GGATGATCTATGATCAGTCAAGAGACATTTGACTGATTCAGATGACACTCGA 2537  
Qy 2532 GGTGCGGCTCAATTTACGATACCTAAACGGGTTGCATTCAAGGTTCTGACCTCTG 2591  
Db 2538 GGCTGCGGCTCATTTACTGTATGCCCCAAGAGACAGCATTCAAAGGTTGCTCCAATC 2597  
Qy 2592 AATCTTCAATGCTCTTCAACCAAGTGTCTGCGTGGGCTCTTGGGTCTATAGACTTTC 2651  
Db 2598 AATCTATCAGACCGTTTGAACCAAGTGTGCGTTGGGCTTTGGGATCAGTAGAAATTTTC 2657  
Qy 2652 TTCAGCAATCATTGCCCTCTTGGTATGGGTATGTTGCGGTCTGAATTTTGGAAAGA 2711  
Db 2658 ATGAGCAGACATTTGCCAATCTGTTATGGCTATGGGGAGAGTCTGAATGCTTGAAGA 2717  
Qy 2712 TTTTCTACATCAATCTCATGCTGTATCTTGGACATCTATTCCTCTTGGCTTACTGT 2771  
Db 2718 TTTGCTATATCAACACCAATGTCTATCTCATTCACCTCTCTTCCACTCATTTGCTATTGC 2777  
Qy 2772 ACATTTGCTGCTCATCTGTTTATTGACAGGGAATTTATCACTCCAGAGCTGAATAATGTT 2831  
Db 2778 ACACTTCAGCGGTGAGTTTGCTCACTGGCAAAATTTGTGATCCCTCAGATCAAGTACTTTT 2837

OY		2832	GCCAGCCTGTGGTTCATGTCACCTTTTATCTGCATTTTTGCTACGAGCATCCTAGAATG	2891
Db		2838	GCAAGTCTATTTTAATAGCTCTTTTCATCTCAATTTTTGCCACTGTATTCTGAAAATG	2897
OY		2892	AGATGGAGTGGTGTGMAATTGATGACTGTGTGAGGAATGACGACTTCTGGGTCAATTGA	2951
Db		2898	AGGTGAGTGGAGTGAGCATTTGAAGATGTGTGGCGAATGACAAGTCTCGGTTATTGGA	2957
OY		2952	GGTGTGCTCACACCCTCTTTGCTGTGTTCCAGGACTCTCAAGTCAATAGCTGTGTT	3011
Db		2958	GGGGTTTCTGCACATTTTTCAGTTATTCBAAGTCTGTCTCAAGTACTGGCAGGCATT	3017
OY		3012	GATACAAGCTTCACCGTGACATCAAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATAT	3071
Db		3018	GATACAATAATTCACAGTCACTGCCAAGGCATCAGATGACGGTAGTTGGGGAAGTGTAT	3077
OY		3072	ACATTCAAATGAGTACCTATTGATACCTCCTACCACCTGTGCTTATTGAAGTTCATT	3131
Db		3078	GCATTCAAATGAGCACCACTCCTCATTCCTCCTACAACCTGTGTCATCAACCTGTG	3137
OY		3132	GGTGTGCTGCTGGCGCTTTCAAAATGCCATCATTAACGGAATATGAGTCATGGGGCCCCCTC	3191
Db		3138	GGGGTGTGTTGGCGTAGCAGATGCAATCAACAATGGAATTCAGTCATGGGGTCTCTC	3197
OY		3192	TTTGGGAAGCTATTCTTTGCAATTTGGGTGATTGTCCATCTTATCCCTTCTCAAAGGT	3251
Db		3198	TTCGGTAAGCTTTTCTTTGCAATCTGGGTCATTGTGCACTGTATCCTTCCTCAAAGGT	3257
OY		3252	TTGTTGGAAGGCAAAAACAGACACCAACGATTGTCAATCGTCTGGTCCATTCTGTGGCT	3311
Db		3258	CTCATGGGCGAGCAGAACCGAACAACCCACCATCGTGTATTGTGTCATTTCTGTGSCA	3317
OY		3312	TCAATCTTCTCGCTCCTTGGGTTCCGATTGATCCTTCTTCCGGAAGGATGATGTCGG	3371
Db		3318	TCTGTTTTCTCTCTTTTCTGGGTAAGAATTGATCCTTCTTGAAGTTAAAGGCCCA	3377
OY		3372	CTTCTTGAGGAGTGTGGTTTGGAATTGC	3398
Db		3378	GATACTAAACAATGTGGCATCAACTGC	3404

RESULT 14	
AAZ99500	
ID	AAZ99500 standard; DNA; 3725 BP.
XX	
AC	AAZ99500;
XX	
DT	03-JUL-2000 (first entry)
XX	
DE	DNA encoding a maize cellulose synthase.
XX	
KW	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW	transgenic plant; plant breeding marker; ss.
XX	
OS	Zea mays.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	179..3398
FT	/*tag= a
FT	/product= "cellulose synthase"
FT	/note= "no termination codon given"
XX	
PN	WO200009706-A2.
XX	
PD	24-FEB-2000.
XX	
PF	16-AUG-1999; 99WO-US18760.
XX	
PR	17-AUG-1998; 98US-0096822.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	

PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
XX  
XX WPI; 2000-224343/19.  
DR P-PSDB; AAY84110.  
XX  
XX  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein -  
XX  
PS Claim 1; Page 105-110; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;

Query Match	35.0%	Score 1378.4;	DB 21;	Length 3725;
Best Local Similarity	66.4%;	Pred. No. 0;		
Matches 2171; Conservative	0;	Mismatches 991;	Indels 108;	Gaps 9;

QY		119	CGCGGGAAGTGCAGGGGAGGAAAGCATGGAAGCCGACGCCCGGGCTGTGGCCGACTCCA	178
Dd		154	CGETGTGTATGAGAGAGCGCGGATGGCGGCCAACAAAGGGATGTGCAGGCTCTCA	213
QY		179	CAACCGCAACGAGCTCGTCGTATCCGCCGACGCGCA-----TCCCCGCCGAAGCC	232
Dd		214	CAACCGCAACGAGTTCTGTATGATCCGCCACGACGCGCCTGTCCCGGCTTAAGCC	273
QY		233	GCCGCGGGAGCAGAACGGGACAGTGTCAGATTGTGCGGACGACGTGGCTTGCCCC	292
Dd		274	CACGAAGAGTGCGAATGGGACAGTGTGCCAGATTGTGCGACACTGTGGGTTTAGC	333
QY		293	CGGCGGGGACCCTCTGCTGCGCTGCACAAGTGGCCTTCCCCTGTGCGGGACTGCTA	352
Dd		334	CACGTGTGATGCTTTGTGCTGCATGAGTGTGCTTCCCTGTCTGCCGCCCTTGCTA	393
QY		353	CGAATACGAGCGCCGGGAGGGCACGCAGAACTGCCCCCAGTGCAGACTTCATACAAGCG	412
Dd		394	TGAGTACGAGCGGCAAGGAAGGGAACCAATGCTGCCCTCAGTGCAGACTAGATACAAGAG	453
QY		413	CCTCAAGGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTCGATGACTTGA	472
Dd		454	ACGAAAGGTAGCCCTCGAGTTCATGTGTGATGATGAGGAGGAAGATGTGATGACTTGA	513
QY		473	CAACGAGTTCAACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTTAACG	532
Dd		514	CAATGAATTCAACTATA-----GCAAGGCAATGGGAAGGG	549
QY		533	CCACATGAGCTACGGCCGTGAGGTGACCTTAATGGCGGCCACAAGCTTTCAGCTCAA	592
Dd		550	CCGAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCATCTGTGCTGCATGACCC	609
QY		593	CCCCAAT--GTTCACTCCTCACCAACGGGCAAAATGTGATGACATCCCAACCGGAGCA	649



Db 610 ACACCATCGGATTCCACGCCCTTACAAGTGAGACAACAGATATCTGAGAGATCCCTGATGC 669  
Qy 650 GCACGCGCTGTGTCCTTCTTTCATGGGTGTGGGGAAAGAGATACATCCCTTCCTTA 709  
Db 670 ATCCCTGACCGTCATTCTATCCGAGT-----CCAACTCGAGCTA 711  
Qy 710 TGGGATCCCACTTACCTGTGCAACCAGTCTATGACCCTCAAGGATCTGTGC 769  
Db 712 TGTGATCCAAAGCTTCCAGTTCGTGAGAGATTGTGACCCCTCGAAAGACTTGAATTC 771  
Qy 770 ATATGGGTATGTAGTGTGCTTGGAAAGAACGGATGAGAAATTGAAAGACAGACAAGA 829  
Db 772 CTATGGGCTTAATAGTGTGACTGGAAGAAAGAGTTGAGAGCTGAGGGGTTAAACAGGA 831  
Qy 830 GAGATGACACAGACGGGAATGATGGTGGTGTGATGATGGTG----- 873  
Db 832 CAAAAATATGTTGCAAGTGACTAATTAATCCAGAGGCTAGAGAGACATGGAGGGGAC 891  
Qy 874 -----ACGATGCTGATCTACCACTAATGAGTGAAGCAAGACAACACTGTCCAGAA 925  
Db 892 TGGCTCAATGGAAGATATGCAAAATGGTGTGATGATGCACGCTTACCTTTGAGCCGCAT 951  
Qy 926 AATTCCACTTCCATCAAGCCAGATTAAATCCATATAGAGATGATTAATTAATTCGGCTTGT 985  
Db 952 TGTGCCAATTTCTCTCAACAGCTCAACCTTTACCGGATAGTAATCATCTCCGCTTAT 1011  
Qy 986 GGTTTGGGGTCTTCTTCCACTACCGAGTATGATCATCCGGTGAATGATGATTTGCTTT 1045  
Db 1012 CATCCTGTGCTTCTTCTCCAAATATCGATCAAGTATCAGTCCAGTGGCTAATGCTATGATTT 1071  
Qy 1046 GTGGCTCATATCTGTATATCTGTGAATCTGGTTTGGCATGTCTTGATTTCTTGATCAATT 1105  
Db 1072 GTGGCTAGTATCTGTATCTGTGAGGTCTGGTTTGCCTTGTCTGGCTTCTAGATCAGTT 1131  
Qy 1106 CCCAAAGTGTTCCTTATTTGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTTCA 1165  
Db 1132 CCCAAATGGTATCCAAATCAACCGTGAGACATATCTGACAGGCTTGCAATTGAGTATGA 1191  
Qy 1166 CAAGGAAGCCAGCATCTCAACTTGTCCAAATTGATTTCTTTGACAGTACGGTTGATCC 1225  
Db 1192 TAGAGAGGAGAGCCATCACAGCTGGCTCCCATTGATGTCTTTGACAGTACAGTGATCC 1251  
Qy 1226 CTTAAGAACTCCTTGTGTCACAAACAATACTGTCTATCTATCTTTCGGTGATTA 1285  
Db 1252 ATTTGAAGAACTCCACTGATCACAGCCAACACTGTTTGTCCATTCTTGTGTGATTA 1311  
Qy 1286 TCCGTGTGATAAGTTTCTTGTCTATGTTTCTGATGATGTTGCTGCAATGCTAACGTTTGA 1345  
Db 1312 CCTGTGTGCAAAAGTGTCTATGTTTCTGATGATGCTCAGCTATGCTGACTTTTGA 1371  
Qy 1346 AGCATATCTGAAACATCTGAATTTGCAAAAGAAATGGGTTCTTCTGCAAAACGTTACAA 1405  
Db 1372 GTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTCCCTTTTGTAAAGACACAA 1431  
Qy 1406 TATTGAACCTCGCGCTCCAGAGTGTACTTCCAAACAGAAAGTACTACTTGAAGACAA 1465  
Db 1432 TATTGAACCAAGAGCTCCAGAAATTTACTTTGCTCAAAAATAGATTACTGAAGACAA 1491  
Qy 1466 GGTGGACGAAACTTTGTAGGAGAGAGACAATGAAGAGAGATATGAGGAATTCAA 1525  
Db 1492 AATTCAACCTTCAATTTGTTAAAGGAAGAGCAATGAAGAGAGATTAAGAAATTCAA 1551  
Qy 1526 GGTGAGATCAATGCTTAGTTGCAAAAGCCAGAAAGTTCTGTAAGAAAGATGAGCAAT 1585  
Db 1552 AATAAGAAATCAATGCCCTTGTGCAAAAGCACAAGAAAGTGTGTAAGAGGGGTGAGCCAT 1611  
Qy 1586 GCAAGATGAACCCCTGTGCTGAAACAATGTTCTGTATCATCTGGAATGATTCAGGT 1645  
Db 1612 GGTGATGGAAGTGTGCTGCAATTAACCTAGGACCATCTGGCATGATTCAGGT 1671  
Qy 1646 CTTCCTTGCCAAAGCGAGGCTTGAAGTGTGAGGGAATGAAGTGCACGATTTGTTTA 1705  
Db 1672 GTTCTTGGGGCAGAGTGTGGCTTGACACTGATGGAATGAATTAACACGCTTGTCTTA 1731

Qy 1706 TGTCTAGAGAAACGACCGCTATTAACCATTAAGAAAGCTGTGCTATGAATGC 1765  
Db 1732 TGTCTCTCGTGAAGAAAGACCGAGCTTTGAGCATCAAGAAAGCTGTGCAATGATGC 1791  
Qy 1766 ATTTGTCGAGTCTCTGTGTAATAAATGCTCATATTTGTTAACTTGAATTGTA 1825  
Db 1792 ACTGATTCGTATCTGTCTGTGCTGACAAATGGTGCCTATCTTCAATGTGATGTGA 1851  
Qy 1826 TCACTACATCAACAACAGCAAGCTATTAAGGAAGCAATGTTTATATGATGACCTTT 1885  
Db 1852 CCAATTACTCAATAGCAGCAAGCTCTTAGAGAAAGCAATGTGCTCATGATGATCCAGC 1911  
Qy 1886 ACTAGAAAGAAAGTTTGTATGTACAGTTCCCTCAAGATTGATGGATGATCCCA 1945  
Db 1912 TCTAGGAAGAAACCTGTTATGTACAATTTCCACAAGATTGATGCAATTGACTTGCA 1971  
Qy 1946 TGACCGATATGCTAACCGGAATGTGTCTTTTGTGATATCAACATGAAGGTTGATGG 2005  
Db 1972 CGATCGATATGCTAATAGAACATAGTCTTCTTGTATCAACATGAAGGCTAGATGG 2031  
Qy 2006 TATTCAGGGTCCAAATTAATGTGTGTAAGTGTGATGATGATTTAGAAAGCAAGCATTAATGG 2065  
Db 2032 CATTCAGGGTCCAGTCTATGTGGGAACAGATGCTGTTCAATAGGCAAGGCTTTGATGG 2091  
Qy 2066 TTAATGATGCCCCCAAAACAAAGAACCCACCATCAAGACTTGCAACTGCTGGCCCAAGTG 2125  
Db 2092 ATATGAT-----CCTGTTTGAATGAGCTGATCTGGAACCTTAACATT 2134  
Qy 2126 GTGCTTTGCTGTGCTGTGCTTGGCAATAGGAAGCAAAAGAAAGTACCAACCAAAAC 2185  
Db 2135 GTTGTAAAG---AGCTGTGTGTGTAAGAAAGAAAGAAACAAGATTATATGATAG 2190  
Qy 2186 AGAGAAAGAAAGTTATTTTCAAGAAAGAAAGAAACCAATCCCTGCATATGCTCT 2245  
Db 2191 TCAAGCCGTA-----TTATGAAGAGAACAGAACTTTCAGCTCCCATCTTAACAT 2241  
Qy 2246 TGTGAAATTGACGAAGCTGCTCCAGAGCTGAGATGAAGAAAGCCGGTATTTGTAATCA 2305  
Db 2242 GGAAGACATCGAGAGGGTATTTGAAGTTATGAGATGAAGGTCAAGTCTATGTCCCA 2301  
Qy 2306 ACAAATAATTAGAAAGAAATTTGGCCAATCTTGTGTTTGTACATCCACACTTCTGCA 2365  
Db 2302 GAGAAATTTGAGAAACGCTTTGTGCTCAGTCCCAATCTTCATTTGATCCACTTTATGAC 2361  
Qy 2366 GAATGTGGAACCTTGAAGGTGCAAGTCCGTCTTCTTTTGAAGAAGCTATACATGT 2425  
Db 2362 TCAAGTGGCATACCACTTCAACAAACCCAGCTTCTACTGAAGGAAGCTATCCATGT 2421  
Qy 2426 CATTAAGTTGTGTTATGAAGCAAGACAGACTGGGGAAGAGATTGGCTGATCTATGG 2485  
Db 2422 TATCAGCTGTGGTACGAGGACAAAACTGAATGGGGAAGAGATTGGCTGATCTATGG 2481  
Qy 2486 ATCAGTTACAGAAATATTTCTAATCTGTTTCAAGATGCAATGTCATGTTGGCGGTCAT 2545  
Db 2482 TTCAGTTACAGAGATATTTCTGACTGGGTTTAAATGCAATGCAAGAGGCTGGCAATCAAT 2541  
Qy 2546 TTAATGATACCTTAACCGGTTGCAATTCAAAGGTTCTGCACTCTGAATCTTTCAAGTCG 2605  
Db 2542 CTAATGATGCCACACGACCTTGTTCAAAGGTTCTGCAACCAATCAATCTTTCTGATCG 2601  
Qy 2606 TCTTCAACAGGTGCTTGGTGGGCTCTTGGTCTATGAGATCTTCTTCAAGCAATGATG 2665  
Db 2602 TCTTAATCAGGTGCTCGTGGGCTCTTGGTCAATGAAATTCGCTTAGCAGACATGG 2661  
Qy 2666 CCTCTTTGGTATGGGTAGGTGGCGGTCTGAATTTTGAAGAAATTTCTCTACATCAA 2725  
Db 2662 TCTTATATGATGCTGCTACATAGGGCGATTGAAGCTTTTGAAGAGGCTGGCTTACATTA 2721  
Qy 2726 CTTCATGCTGATCTTGGACATCTATTTCCCTCTTGGCTTAATGATGATGCTGCCAT 2785  
Db 2722 CACCATGTTTATCCAAATCAATCTGTTCCGCTTATCGCCTATTTGTGTCTTCTGCTAT 2781

OY	2786	CTGTTTATTGACAGGGAAATTTATTCACCTCCAGAGCCTGAATAATGTGCCAGCCTTGCGTT	2845
Db	2782	CTGTCTTCTTACCATAAATTAATTCATTCTCGAGATTAGTAATTATGCTGGAATGTTCTT	2841
OY	2846	CATGTCACCTTTTATCTGCATTTTGTCTACGAGCATCTGAATAATGAGATGAGTGGTGT	2905
Db	2842	CATTCTTCTTTTGGCCCTCCATTTTTCGCAACTGTATATTGAGCTCAGATGAGTGGTGT	2901
OY	2906	TGGAATTGATGACTGGTGGAGAATGAGCAGTCTGGGTCAATTGGAGGTGTGTCCTACA	2965
Db	2902	TGGCATTGAGATTTGGTGGAGAAATGAGCAGTTTTGGGTATTGTGTGGCACCTCTGCCCA	2961
OY	2966	CCTCTTTGCTGTGTTCCAGGCACTTCTCAAGTCATAGCTGTGTTGATACAAGCTTCAC	3025
Db	2962	TCTCTTCCGGGTGTTCCAGGGTCTGCTGAAGTGTTGGCTGGGATTGATACCAACTTCAC	3021
OY	3026	CGTGACATCAAAGG--GTGGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATG	3082
Db	3022	AGTTACCTCAAAGGCATCTGATGAGGATGGCGCACTTGTGAGCTATATGTTCAGATG	3081
OY	3083	GACTACCTTATTGATACCTCCACCACCTTGCTTCTATTGAACCTTCATTGTGTGTGCGC	3142
Db	3082	GACCAAGTTTGCTCATCCCTCCGACCACCTGTTCTGTGCATTAACCTGTCGGAATGTTGCGC	3141
OY	3143	TGGCGTTTCCAATGCGATCAATAACGATATGAGTCATGGGGCCCCCTCTTGGGAAGCT	3202
Db	3142	AGGAATTTGATATGCCATTAAACAGCGGCTAACCAATCCTGGGGTCCGCTCTTGGAAAGCT	3201
OY	3203	AATCTTTGCATTTTGGGTGATGTCATCTTATCCCTTCTCAAAAGTTTGGTTGGAG	3262
Db	3202	GTTCTTCTCGATCTGGGTGATCTCCATCTACCCCTCTCTCAAGGGTCTCATGGGAG	3261
OY	3263	GCAAAACAGGACACCAACGATTTGTCATGCTGTGTCATCTGTGCTTCAATCTTCTC	3322
Db	3262	GCAGAACCGCAGCGCCAACAAATCGTCATGTTGTGTCATCCTCTTGGCGTATCTTCTC	3321
OY	3323	GCTCCTTTGGGTTCCGATTCCTTTCT	3352
Db	3322	CTTGCTGTGGGTGAAGATCGATCTTTTCAT	3351

RESULT	15
AAZ99506	
ID	AAZ99506 standard; DNA; 3725 BP.
XX	
AC	AAZ99506;
XX	
DT	03-JUL-2000 (first entry)
XX	
DE	DNA encoding a maize cellulose synthase.
XX	
KW	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW	transgenic plant; plant breeding marker; ss.
XX	
OS	Zea mays.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	179..3397
FT	/tag= a
FT	/product= "cellulose synthase"
XX	/note= "no termination codon given"
PN	WO200009706-A2.
XX	
PD	24-FEB-2000.
XX	
PF	16-AUG-1999; 99WO-US18760.
XX	
PR	17-AUG-1998; 98US-0096822.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Dhuga KS, Helentjaris TG, Bowen BA, Wang X;

XX WP1; 2000-224343/19.  
 DR P-PSDB; MAY84112.  
 XX  
 PT New genes which encode maize cellulose synthase polypeptides in plants  
 PT useful for modulating the expression of cellulose synthase in plants  
 PT and to produce transgenic plants expressing the novel protein  
 XX  
 PS Claim 1; Page 121-126; 119pp; English.  
 XX  
 CC The present sequence encodes a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain  
 CC antibodies specifically immunoreactive with a protein.  
 CC  
 XX  
 XX Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;  
 SQ  
 Query Match 35.0%; Score 1378.4; DB 21; Length 3725;  
 Best Local Similarity 66.4%; Pred. No. 0;  
 Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9

QY	119	CGCGCGAAGTGGAGGGGAGGAAGCGATGAGGCGCGCGCTGTGGCCGCTCCA	178
DB	154	CGGTGTGTGTAGAGGAGCGCGGAGATGGCGCCAAAGGGAATGTGGCAGGCTCTCA	213
QY	179	CAACCGCAACGAGCTCGTGTCTCATCCGCCGCGAGCGCGA-----TCCCGGGCCGAAGCC	232
DB	214	CAACCGCAACGAGTTCGTATGATCCGCCACGACGCGAGCGCGCTGTCCCGGCTAAGCC	273
QY	233	GCGCGGGAGCAGAACGGGCGAGTGTGCCAGATTGTGGCGCAGACGTGGCCTTGCCCC	292
DB	274	CACGAGAAGTGGCAATGGGCAAGTCTGCCAGATTGTGTGGCACACTGTGGCGTTCAAGC	333
QY	293	CGCGCGGAGCCCCCTTCGTGGCGGTCAACGAGTGGCGCTTCCCGCTTGCCGGGACTGCTA	352
DB	334	CACGTGTGATGTCTTGTGTGCTTGCATGAGTGTGCCCTTCCCTGTGCGCCCTTGCTA	393
QY	353	CGAATACGAGCGCGCGGAGGGGACGACGAAGTGCCTCCCGAGTGCAGACTGATACAAAGCG	412
DB	394	TGAGTACGAGCGCAAGGAAGGAACCAATGTCTGCCCTCAGTGCAGACTAGATCAAGAG	453
QY	413	CCTCAAGGCTGCCAACGTGTGACCGGTGACGAGGAGGAGAGCGCGCTCGATGACCTGGA	472
DB	454	ACAGAAAGGTAGCCCTCGAGTTCAATGTGATGATGAGAGGAAGATGTGATGACCTGGA	513
QY	473	CAACGAGTTCACTGGGACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTCTACGG	532
DB	514	CAATGATTTCACTATA-----GCAAGGCAATGGGAAGG	549
QY	533	CCACATGAGCTACGGCGGTGAGGTGACCCCTAATGCGCGCCACAAGCTTTCCAGCTCAA	592
DB	550	CCGAGAGTGGCAGCTTCAAGGAGATGACGCTGATCTGTCTTCATCTGCTCGCCATGACCC	609
QY	593	CCCCAAT---GTTCCACTCTCACCAACGGGCAAAATGTTGTGATGATATCCACCGGAGCA	649
DB	610	ACACATTCGATTCACGCGCTTACAGTGAAGTGAACAACAGATATCTGAGAGAGATCCCTGATGC	669

QY	650	GCACGCGTGTGCTCTTTCATGCGTGTGGGGAAGAGATACATCCCTTCTTA	709
Db	670	ATCCCTGACCGTCATTCTATCCGAGT-----CCAACTGAGCTA	711
QY	710	TGCGATCCAGCTTACTGTGCAACCAGTCTATGACCATCCAAGATCTTGCTGC	769
Db	712	TGTTGATCCAGCGTTCAGTTCCTGTGAGGATGTGACCCCTCGAAGACTGAATTC	771
QY	770	ATATGGTATGTAGTGTGCTTGGAAGAACGGATGGAATTGGAAGACAGACAGA	829
Db	772	CTATGGCTTAATAGTGTGACTGGAAGAAAAGAGTTGAGAGCTGAGGGTTAAACAGA	831
QY	830	GAGGATGCACCAGACGGGAAATGATGCTGTGTGATGATGTTGTTG-----	873
Db	832	CAAAATATGTTGCAAGTGACTAATAATATCCAGAGCTAGAGAGACATGAGGGAC	891
QY	874	-----ACGATGCTGATCTACCACTAATGATGAAGCAACAACACTGTCCAGAA	925
Db	892	TGGCTCAATGAGAGAATATGCAAAATGTTGATGATGCACGCCCTACCTTGAGCCGAT	951
QY	926	AATTCCACTTCCATCAAGCCAGATTAATCCATATAGATGATTAATCATTTATCCGCTTGT	985
Db	952	TGTGCCAATTTCTCAACCACTCAACCTTACCGGATAGTAATCATTTCTCCGCTTAAT	1011
QY	986	GGTTTGGGGTCTTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGATTTGCTTT	1045
Db	1012	CATCTGTGCTTCTTCCAAATATCGTATCAGTCATCCAGTCGTAATGCTTATGATTT	1071
QY	1046	GTGCTCATATCTGTATCTGTAAATCTGTTGCCATGCTTGATTCATTAAT	1105
Db	1072	GTGCTAGTATCTGTATCTGTAGAGTCTGTTGCCCTTCTGCTTCTAGATCAGTT	1131
QY	1106	CCCAAAGTGTCCCTATGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTGCA	1165
Db	1132	CCCAAATGATATCCAAATCAACCGTAGACATATCTGCAGAGCTTGCAATGAGGTATGA	1191
QY	1166	CAAGGAAGGCCAGCCATCTCACTTGCTCCAATTGATTTCTTGTCACTACGTTGATCC	1225
Db	1192	TAGAGAGGAGAGCCATCACTGAGCTGCCATTTGATGTCTTGTCACTACGTTGATCC	1251
QY	1226	CTTAAAGGAACCTCCTTGGTCAACAACAATACTGTTCTATCTATCCTTTCGGTGAATTA	1285
Db	1252	ATTGAAGAACTCCACTGATCAACGCCAACACTGTTGTTCATCTTGTGTGATTA	1311
QY	1286	TCCGTGTGATTAAGTTTCTGCTATGTTTCTGATGATGCTGCAATGCTAACGTTTGA	1345
Db	1312	CCCTGTTGACAAAGTGTCACTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTGA	1371
QY	1346	AGCATTATCTGAACATCTGAATTTGCCAAGAAATGGGTTCTTCTGCAACCGTACAA	1405
Db	1372	GTTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGTTCCCTTTGTAAAGACACAA	1431
QY	1406	TATTGAACCTCGCGCTCCAGAGTGTACTTCCAACAGAAATAGACTACTTGAAGACAA	1465
Db	1432	TATTGAACCAAGAGCTCCAGAAATTTACTTGTCTCAAAAAATAGATTACCTGAAGCAAA	1491
QY	1466	GGTGGCAGCAAACTTGTTAGGAGAGAGAGCAATGAGAGAGATGAGAAATCAA	1525
Db	1492	AATTCAACCTTCATTTGTTAAGAAAAGACGAGCAATGAGAGATGAGAAATCAA	1551
QY	1526	GGTGAATCAATGCTTGTAGTTCCAAAGCCAGAAAGTTCTGAAGAAGATGACAAT	1585
Db	1552	AATGAATCAATGCTTGTGTTGCCAAAGCACAGAAAGTCTGAAAGAGGTGACCAT	1611
QY	1586	GCAAGATGAACCCCTGGCTGGAACAATGTTCTGATCATCTGGAATGATTCAGGT	1645
Db	1612	GGCTGATGGAATGCTTGGCTTGGGAATTAACCTTAGGACCATCTTGCAATTCAGGT	1671
QY	1646	CTTCTTGGCCAAAGCGAGGCTTGACTGTGAGGGAATGAATGCCACGATTTGTTTA	1705
Db	1672	GTTCTTGGGGCACAGTGTGGGCTTGACACTGATGGAATGAATTAACACGCTTGTGCTA	1731

QY	1706	TGTTTCTAGAGAAA	CGACCAGGCTATTAACCATCATAGA	AAAAGCTGGTCTAGTAATGC	1765	
Db	1732	TGTCCTCGTGAAGA	AGACCAGGCTTTCAGCATCAACA	GAAGGCTGGTGCATGAATGC	1791	
QY	1766	ATTGTCGAGTCTCTG	CTGTACTAACAAATGCTCCATATTTG	TAACTTGAATGTGA	1825	
Db	1792	ACTGATTCGTATCTG	CTGTGCTGACCAATGGTGCCATCTTCT	CAATGTGATGTGA	1851	
QY	1826	TCACTACATCAACA	CAAGCAAGGCTATAAAGAACCAATGTG	TTTATGATGACCCTTT	1885	
Db	1852	CCATTACTTCAATAG	CAGCAAAAGCTCTTAGAAGAACCAATGTG	CTTCATGATGATCCAGC	1911	
QY	1886	ACTAGGAAGAAGTTT	GCTATGTACAGTTCCTCAAAAGATTG	ATGGGATGATCGCCA	1945	
Db	1912	TCTAGGAAGGAAC	TTGTTATGTACAATTTCCACAAAGATTG	ATGGCATTGACTTGCA	1971	
QY	1946	TGACCGATATGCTAA	CCGGAATGTGCTTTTGTATATCAACATGA	AAAGTTTGGATGG	2005	
Db	1972	CGATCGATATGCTAA	TAGAACATAGTCTTTTGATATCAACATGA	AAAGGTTAGATGG	2031	
QY	2006	TATTCAGGGTCCAA	TTTATGTTGGTACTGATGTATTTAGAAGG	CAGGCATTAATAGG	2065	
Db	2032	CATTCAAGGTCCAGT	CTATGTGGAAACAGATGCTGTTCAATAGG	CAGGCTTTGTATGG	2091	
QY	2066	TTATGATGCCCCAA	AAACAAGAACCAATCAAGGACTTGCAACTG	CTGGCCAAATG	2125	
Db	2092	ATATGAT-----	-----CCTGTTTGA	CTGAAGCTGATCTG	GAACCTAACATT	2134
QY	2126	GTGCTTTTGCTGTG	CTGCTTTGGCAATAGAACAAAGAAGACTA	CCAAACCCAAAC	2185	
Db	2135	GTTGTTAG-----	AGCTGCTGTGCTAG	AAGAACAAAGAACAAGAGTTATATG	ATAG	2190
QY	2186	AGAGAAGAAAAAG	TATATTTTCAAGAAAGAACAAACCAATCCC	CTGATATGCTCT	2245	
Db	2191	TCAAAGCCGTA-----	TTATGAAGAGAACAAATCTTCA	GCTCCATCTTTAACAT	2241	
QY	2246	TGCTGAATTTGACGA	AGCTGCTCCAGAGCTGAGAAATGAAAGG	CCCGTATTTGTAATCA	2305	
Db	2242	GGAAGACATCGAG	AGGTATTTGAAGTATGAGGATGAAGG	TCAGTCTTATGTCCA	2301	
QY	2306	ACAAAAATTAGAAA	GAATTTGGCCAATCTTCTGTTTGTACATCC	ACACTTCTCGA	2365	
Db	2302	GAGGAATTTGAGAA	ACGCTTTGTCAGTCTCCAATCTTCATTG	CATCCACTTTATGAC	2361	
QY	2366	GAATGCTGAACCTT	GGAAGTGCAGAGTCTCTCTTTTGAAGAAG	CTATACATGT	2425	
Db	2362	TCAAGGTGCAATAC	CACTTCAACAAACCAAGTCTCTACTGAAG	GAAGCTATCCATGT	2421	
QY	2426	CATTAGTTGTGTTAT	GAAAGCAAGACAGACTGGGAAAAAGATTTG	CGCTGATCTATGG	2485	
Db	2422	TATCAGCTGTGGT	ACGAGACAAAAGTGAATGGGAAAAAGATTTG	CGATCTATGG	2481	
QY	2486	ATCAGTTACAGAA	AGATATTTCTAATCTGTTTCAAGATGCAT	TGTCATGTTGGCGCTCAAT	2545	
Db	2482	TTCAAGTTACAGAG	AGATATTTCTGACTGGGTTTAAATGCATG	CAAGAGGCTGGCAATCAAT	2541	
QY	2546	TTACTGCATACCT	TAAAACGGGTGCATTTCAAAGGTTCTGCA	CCCTGAAATCTTTCAGATCG	2605	
Db	2542	CTACTGCATGCCA	CCAGCAAGCTTGTTCAAAGGTTCTGCA	CCCAATCAATCTTTCGATCG	2601	
QY	2606	TCTTCAACAGTGCT	TCGGGTGGGCTCTTGGGTCTATGAGATCT	TTCTTCAGCAATCATG	2665	
Db	2602	TCTTAATCAGTGCT	CCGTGGGTGGGCTCTTGGGTCAGTGAA	AATTCGCTTAGCAGACATGG	2661	
QY	2666	CCCTCTTTGGTATG	GGGTATGTTGGCGGTCTGAAATTTT	TGAAAAGATTTTCTACATCAA	2725	
Db	2662	TCCTATATGCTATG	GGCTACAATGGCGCATTGAAGCTTTT	TGAGAGGCTGGCTTACATTA	2721	
QY	2726	CTCCATCGTATTC	CTTGGAACATCTATTCCTCTTGGCTTACTG	TACATTTGCTGCCAT	2785	
Db	2722	CACCATTTGTTAT	CCAATCACATCTGTTCCGCTTATCG	CCCTAATTGTCCTGCTAT	2781	
QY	2786	CTGTTTATTTGAC	GGAATTTTATCACTCCAGAGCTGAATATG	TTTGGCAGCCTGTGCTT	2845	



Db	2782	CTGCTCTTACCAATAAATTATCATCTCCTGAGATTAGTAATTATGCTGGAACTGTTCTT	2841
QY	2846	CATGTCACCTTTTATCTGCATTTTCTACGAGCATCTAGAAATGAGATGAGTGTGT	2905
Db	2842	CATTCTCTTTTGGCCTCCATTTTCGCAACTGGTATATTGGAAGCTCAGATGAGTGTGT	2901
QY	2906	TGGAATTGATGACTGGTGGAGGAATGAGCAGTTCTGGGTCAATTGGAGGTGTCTCACA	2965
Db	2902	TGGCATTTGAGATTGGTGGAGAAATGAGCAGTTTGGGTATTGGTGGCACCCTTGCCCA	2961
QY	2966	CCTCTTGTGCTGTTCACGGAACCTTCTCAAGTCATAGCTGTGTGATACAAAGCTTAC	3025
Db	2962	TCTCTTGGCGGTGTTCAGGGTCTGCTGAAAGTGTGGCTGGGATTGATACCAACTTAC	3021
QY	3026	CGTGACATCAAAG--GTGGAGATGATGAGAGTTCTCAGAGCTATATACATTCAAATG	3082
Db	3022	AGTTACCTCAAGGCATCTGATGAGGATGGCGACTTGGCTGAGCTATATGTGTTCAAGTG	3081
QY	3083	GACTACCTTATGATATCCTCTTACCACCTTGCTTCTATTGAACTTCATTGCTGTGTGCGC	3142
Db	3082	GACCAGTTTGTCTCATCCCTCCGACCACACTGTTCTTGTCATTAACTGGTCGGAATGTGGC	3141
QY	3143	TGGCGTTTCAAAATGCGATCAATAACGAGATATGATCATGGGGCCCCCTTTGGGAAGCT	3202
Db	3142	AGGAATTTGATGCGCATTAACAGCGGCTACCAATCTGGGGTCCGCTCTTTGGAAGCT	3201
QY	3203	ATTCTTTGCATTTTGGGTGAATGTCCATCTTTATCCCTTCTCAAGGTTTGGTTGGAAG	3262
Db	3202	GTTCTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCTCAAGGTCATGGGAG	3261
QY	3263	GCAAAACAGGACACCAACGATTTGTCATGCTGTGTCATTCGTGCTTCAATCTTCTC	3322
Db	3262	GCAGAACCGCACGCCCAACAATGTCATCGTTTGGTCCATCCTCCTTGGCTATCTTCTC	3321
QY	3323	GCTCCTTTGGGTTGGATTGATCTTTCT 3352	
Db	3322	CTTGTCTGTGGTGAAGATGATCTTTCTAT 3351	

Search completed: December 15, 2003, 18:48:39  
Job time : 979 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 22:42:15 ; Search time 1271 Seconds  
(without alignments)  
10292.436 Million cell updates/sec

Title: US-09-720-383C-9  
Perfect score: 3936  
Sequence: 1 cttccctcgtcgtcgtcg...ctgagtggtttgtgaaaa 3936

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3936	100.0	3936	10	US-09-900-237-9 Sequence 9, Appl1
2	3936	100.0	3969	13	US-10-209-059-49 Sequence 49, Appl1
3	3936	100.0	3969	13	US-10-160-719-17 Sequence 17, Appl1
4	3936	100.0	3969	13	US-10-160-719-37 Sequence 37, Appl1
5	2592.6	65.9	3776	10	US-09-900-237-3 Sequence 3, Appl1
6	2546.8	64.7	3568	13	US-10-209-059-41 Sequence 41, Appl1
7	2546.8	64.7	3568	13	US-10-160-719-1 Sequence 1, Appl1
8	2241	56.9	3786	10	US-09-900-237-7 Sequence 7, Appl1
9	2241	56.9	3813	13	US-10-209-059-17 Sequence 17, Appl1
10	2241	56.9	3813	13	US-10-160-719-25 Sequence 25, Appl1
11	2241	56.9	3813	13	US-10-160-719-45 Sequence 45, Appl1
12	1509	38.3	3255	10	US-09-938-842A-2530 Sequence 2530, App
13	1431.8	36.4	3255	10	US-09-938-842A-559 Sequence 559, App
14	1429.4	36.3	3828	15	US-10-229-193-7 Sequence 7, Appl1
15	1378.4	35.0	3725	13	US-10-209-059-45 Sequence 45, Appl1

16	1378.4	35.0	3725	13	US-10-160-719-13	Sequence 13, Appl1
17	1378.4	35.0	3725	13	US-10-160-719-21	Sequence 21, Appl1
18	1378.4	35.0	3725	13	US-10-160-719-41	Sequence 41, Appl1
19	1365	34.7	3753	13	US-10-160-719-33	Sequence 33, Appl1
20	1365	34.7	3753	13	US-10-160-719-53	Sequence 53, Appl1
21	1365	34.7	3780	13	US-10-209-059-1	Sequence 1, Appl1
22	1365	34.7	3780	13	US-10-160-719-9	Sequence 9, Appl1
23	1331.8	33.8	2890	10	US-09-900-237-17	Sequence 17, Appl1
24	1330.8	33.8	3673	15	US-10-229-193-11	Sequence 11, Appl1
25	1326.6	33.7	3603	15	US-10-229-193-5	Sequence 5, Appl1
26	1322.4	33.6	3704	13	US-10-209-059-13	Sequence 13, Appl1
27	1322.4	33.6	3704	13	US-10-160-719-57	Sequence 57, Appl1
28	1288.2	32.7	3746	13	US-10-160-719-29	Sequence 29, Appl1
29	1288.2	32.7	3746	13	US-10-160-719-49	Sequence 49, Appl1
30	1288.2	32.7	3773	13	US-10-209-059-9	Sequence 9, Appl1
31	1288.2	32.7	3773	13	US-10-160-719-5	Sequence 5, Appl1
32	1287.2	32.7	3563	10	US-09-900-237-25	Sequence 25, Appl1
33	1279.4	32.5	3799	13	US-10-267-459-5	Sequence 5, Appl1
34	1279.4	32.5	3799	13	US-10-209-059-21	Sequence 21, Appl1
35	1271	32.3	3626	10	US-09-900-237-29	Sequence 29, Appl1
36	1269.2	32.2	3614	15	US-10-229-193-9	Sequence 9, Appl1
37	1254.2	31.9	3517	10	US-09-900-237-13	Sequence 13, Appl1
38	1218	30.9	2663	10	US-09-900-237-23	Sequence 23, Appl1
39	1170	29.7	2830	13	US-10-267-459-1	Sequence 1, Appl1
40	1170	29.7	2830	13	US-10-209-059-5	Sequence 5, Appl1
41	1138.8	28.9	3231	13	US-10-209-059-27	Sequence 27, Appl1
42	1134.8	28.8	3328	10	US-09-838-539-1	Sequence 1, Appl1
43	1016.6	25.8	3028	13	US-10-209-059-29	Sequence 29, Appl1
44	1004	25.5	3470	13	US-10-209-059-25	Sequence 25, Appl1
45	884.6	22.5	2248	15	US-10-229-193-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1  
US-09-900-237-9  
; Sequence 9, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 3936  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-900-237-9

QY	1	CTTCTCCCTCGTGGGTGCGGCGGCTCGGCGCTTGGTGAGAAACCACTCGGGG	60
DB	1	CTTCTCCCTCGTGGGTGCGGCGGCTCGGCGCTTGGTGAGAAACCACTCGGGG	60
QY	61	ATGAGATCTGCTGCTAGAGTGAAGAGCTACCGTCACTTCTGCTTCTGCTGCGG	120
DB	61	ATGAGATCTGCTGCTAGAGTGAAGAGCTACCGTCACTTCTGCTTCTGCTGCGG	120
QY	121	GCGGAAGTGAGGGGAGGAGGATGAGGGGAGCGGCGGCTGTGGCGGCTCCACA	180

Db 121 GCGAAGTGGAGGGAGAAAGCATGGAGGCGAGCGCCGGGCTGTGGCCGGCTCCACA 180  
QY 181 ACCGCAACGAGCTCGTCGTATCCGCCCGGACGCGCATCCGGGCGGAAGCCGCGGG 240  
Db 181 ACCGCAACGAGCTCGTCGTATCCGCCCGGACGCGCATCCGGGCGGAAGCCGCGGG 240  
QY 241 AGCAGAACGGGCAAGGTGTGCCAGATTTCGGCGACGACGTGGCCTTGCCCCGGCGGG 300  
Db 241 AGCAGAACGGGCAAGGTGTGCCAGATTTCGGCGCGACGACGTGGCCTTGCCCCGGCGGG 300  
QY 301 ACCCTTCTGTGGCGTGCAACGAGTGCCTTCCCCGTCTGCCGGGACTGTCTACGAATACG 360  
Db 301 ACCCTTCTGTGGCGTGCAACGAGTGCCTTCCCCGTCTGCCGGGACTGTCTACGAATACG 360  
QY 361 AGCGCCGGGAGGGCACGCAAGTCCCCAGTGCAAGACTCGATACAAAGCCCTCAAGG 420  
Db 361 AGCGCCGGGAGGGCACGCAAGTCCCCAGTGCAAGACTCGATACAAAGCCCTCAAGG 420  
QY 421 GCTGCCAACGTGTGACCCGGTGAAGAGAGAGAGACGGCGTCGATGACCTGACAAAGAGT 480  
Db 421 GCTGCCAACGTGTGACCCGGTGAAGAGAGAGAGACGGCGTCGATGACCTGACAAAGAGT 480  
QY 481 TCAACTGGGACGGCCATGACTCCGACTCTGTGCCGAGTCCATGCTCTACGGCCACATGA 540  
Db 481 TCAACTGGGACGGCCATGACTCCGACTCTGTGCCGAGTCCATGCTCTACGGCCACATGA 540  
QY 541 GCTACGGCCGTGGAGGTGACCCCTAATGGCGCGCACAAAGCTTTCAGCTCAACCCCAATG 600  
Db 541 GCTACGGCCGTGGAGGTGACCCCTAATGGCGCGCACAAAGCTTTCAGCTCAACCCCAATG 600  
QY 601 TTCACCTCTCACCACGCGGCAATGGTGATGACATCCACCGGAGCAGCAGCGCGTGG 660  
Db 601 TTCACCTCTCACCACGCGGCAATGGTGATGACATCCACCGGAGCAGCAGCGCGTGG 660  
QY 661 TGCCCTCTTTCATGGGTGGTGGGGAAGAAGATACATCCCTTCTTATGCGGATCCCA 720  
Db 661 TGCCCTCTTTCATGGGTGGTGGGGAAGAAGATACATCCCTTCTTATGCGGATCCCA 720  
QY 721 GCTTACCTGTGCAACCCAGTCTATGGAACCATCCAGAGTCTTGCTGCATATGGGTATG 780  
Db 721 GCTTACCTGTGCAACCCAGTCTATGGAACCATCCAGAGTCTTGCTGCATATGGGTATG 780  
QY 781 GTAGTGTGCTTGGAAAGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGATGACCC 840  
Db 781 GTAGTGTGCTTGGAAAGAACGGATGGAGAATTGGAAGCAGAGACAAGAGAGATGACCC 840  
QY 841 AGACGGGGAATGATGGTGGTGGTGAATGGTGACGATGCTATCCACTAATGATG 900  
Db 841 AGACGGGGAATGATGGTGGTGGTGAATGGTGACGATGCTATCCACTAATGATG 900  
QY 901 AAGCAAGACAACAACCTGTCCAGGAAATTCACCTTCATCAAGCCAGATTATCCATATA 960  
Db 901 AAGCAAGACAACAACCTGTCCAGGAAATTCACCTTCATCAAGCCAGATTATCCATATA 960  
QY 961 GGATGATATCATATTTCGGCTGTGGTTTGGGGTCTTCTTCCACTACCGAGTATGC 1020  
Db 961 GGATGATATCATATTTCGGCTGTGGTTTGGGGTCTTCTTCCACTACCGAGTATGC 1020  
QY 1021 ATCCGGAATGATGATTTGCTTTGTGGCTCATATCTGTTAATCTGTGAATCTGGTTG 1080  
Db 1021 ATCCGGAATGATGATTTGCTTTGTGGCTCATATCTGTTAATCTGTGAATCTGGTTG 1080  
QY 1081 CCATGTCTTGGAATCTTGATCAATTCCAAAGTGTTCCCTATGAGAGAGACTTAAC 1140  
Db 1081 CCATGTCTTGGAATCTTGATCAATTCCAAAGTGTTCCCTATGAGAGAGACTTAAC 1140  
QY 1141 TAGACGGGCTGTCACTGAGGTTGCAAGGAAGCCAGCACTCAACTGTCTCAATTG 1200  
Db 1141 TAGACGGGCTGTCACTGAGGTTGCAAGGAAGCCAGCACTCAACTGTCTCAATTG 1200  
QY 1201 ATTTCTTTGTCAGTACGTTGATCCCTTAAAGAACCTCTTGGTCAACAACAATACTG 1260  
Db 1201 ATTTCTTTGTCAGTACGTTGATCCCTTAAAGAACCTCTTGGTCAACAACAATACTG 1260

QY 1261 TTCTATCTATCCCTTTCGGTGGAATTATCCTGTGTGATAAGGTTCTTGCTATGTTCTGATG 1320  
Db 1261 TTCTATCTATCCCTTTCGGTGGAATTATCCTGTGTGATAAGGTTCTTGCTATGTTCTGATG 1320  
QY 1321 ATGTGCTGCAATGCTAAACGTTTGAAGCATTAATCTGAAAACATCTGAATTGCAAGAAT 1380  
Db 1321 ATGTGCTGCAATGCTAAACGTTTGAAGCATTAATCTGAAAACATCTGAATTGCAAGAAT 1380  
QY 1381 GGGTCCCTTCTGCAACCGGTACAATATGAACTTCGCGCTCCAGAGTGTACTTCCAAC 1440  
Db 1381 GGGTCCCTTCTGCAACCGGTACAATATGAACTTCGCGCTCCAGAGTGTACTTCCAAC 1440  
QY 1441 AGAAGATAGACTACTTGAAGAACAAGTGGCAGCAACTTTGTTAGGAGAGAGAGACAA 1500  
Db 1441 AGAAGATAGACTACTTGAAGAACAAGTGGCAGCAACTTTGTTAGGAGAGAGAGACAA 1500  
QY 1501 TGAAGAGAGATGAGGAATTCAAGGTGAGAAATCAATGCCCTTAGTTGCCAAAGCCAGA 1560  
Db 1501 TGAAGAGAGATGAGGAATTCAAGGTGAGAAATCAATGCCCTTAGTTGCCAAAGCCAGA 1560  
QY 1561 AAGTCCCTGAAGAGGATGACCAATGCAAGATGGAACCCCTGCGCTGGAACAATGTTTC 1620  
Db 1561 AAGTCCCTGAAGAGGATGACCAATGCAAGATGGAACCCCTGCGCTGGAACAATGTTTC 1620  
QY 1621 GTGATCATCCTGGAATGATTCAGGTCTTCTGCGCAAGCGGAGCCTGACTGTGAGG 1680  
Db 1621 GTGATCATCCTGGAATGATTCAGGTCTTCTGCGCAAGCGGAGCCTGACTGTGAGG 1680  
QY 1681 GAAATGAACCTGCCACGATTGGTTTATGTTCTAGAGAGAAACGACGAGCTATAACCATC 1740  
Db 1681 GAAATGAACCTGCCACGATTGGTTTATGTTCTAGAGAGAAACGACGAGCTATAACCATC 1740  
QY 1741 ATAAGAAAGCTGTGCTATGAAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800  
Db 1741 ATAAGAAAGCTGTGCTATGAAATGCATTGGTCCGAGTCTCTGCTGTACTAACAAATGCTC 1800  
QY 1801 CATATTTGTTAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860  
Db 1801 CATATTTGTTAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGGAAG 1860  
QY 1861 CAATGTGTTTATGATGGAACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTTCCTC 1920  
Db 1861 CAATGTGTTTATGATGGAACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTTCCTC 1920  
QY 1921 AAAGATTTGATGGGATGATCGCCATGACCGATATGCTAACCCGAATGTGCTTTTTTG 1980  
Db 1921 AAAGATTTGATGGGATGATCGCCATGACCGATATGCTAACCCGAATGTGCTTTTTTG 1980  
QY 1981 ATATCAACATGAAGGTTTGATGATGTATTCAGGGTCCAATTTATGTTGTACTGATGTG 2040  
Db 1981 ATATCAACATGAAGGTTTGATGATGTATTCAGGGTCCAATTTATGTTGTACTGATGTG 2040  
QY 2041 TATTTAGAAAGGCAAGCATTAATGTTATGATGCCCCCAAAACAAGAACCCATCAA 2100  
Db 2041 TATTTAGAAAGGCAAGCATTAATGTTATGATGCCCCCAAAACAAGAACCCATCAA 2100  
QY 2101 GGACTTGCAACTGTGCCCCAAGTGTGCTTTGCTGTGCTTGGCAATAGGAAGC 2160  
Db 2101 GGACTTGCAACTGTGCCCCAAGTGTGCTTTGCTGTGCTTGGCAATAGGAAGC 2160  
QY 2161 AAAAGAGACTACCAAAACCAAAACAGAGAGAAAAAGTTATATTTCAGAAAAAGAG 2220  
Db 2161 AAAAGAGACTACCAAAACCAAAACAGAGAGAAAAAGTTATATTTCAGAAAAAGAG 2220  
QY 2221 AGAACCAATCCCTGCAATATGCTCTTGTGTAATTGACGAAGCTGCTCCAGAGCTGAGA 2280  
Db 2221 AGAACCAATCCCTGCAATATGCTCTTGTGTAATTGACGAAGCTGCTCCAGAGCTGAGA 2280  
QY 2281 ATGAAAAAGCCGGTATTTGTAATCAACAATAATTAGAAAAAGAAATTTGGCCAATCTTCTG 2340  
Db 2281 ATGAAAAAGCCGGTATTTGTAATCAACAATAATTAGAAAAAGAAATTTGGCCAATCTTCTG 2340



QY 2341 TTTTGTACATCCACACTTCTCGAAGATGTTGGAACCTTGAAAGTGCAGTCTGCTT 2400  
|||||  
Db 2341 TTTTGTACATCCACACTTCTCGAAGATGTTGGAACCTTGAAAGTGCAGTCTGCTT 2400  
QY 2401 CTCTTTGAAAGAGCTATACATGTCATAGTTGTGTATGAAAGACAAGACAGACTGGG 2460  
|||||  
Db 2401 CTCTTTGAAAGAGCTATACATGTCATAGTTGTGTATGAAAGACAAGACAGACTGGG 2460  
QY 2461 GAAAGAGATTGGCTGATCTATGATCAGTTACAGAGATATTTCTACTGTTTCAAGA 2520  
|||||  
Db 2461 GAAAGAGATTGGCTGATCTATGATCAGTTACAGAGATATTTCTACTGTTTCAAGA 2520  
QY 2521 TGCATTGTTCATGTTGGCGGTCAATTTACTGCATACCTTAAACGGTTGCATTCAAGGTT 2580  
|||||  
Db 2521 TGCATTGTTCATGTTGGCGGTCAATTTACTGCATACCTTAAACGGTTGCATTCAAGGTT 2580  
QY 2581 CTGCACCTCTGAATCTTTACAGATCGTCTTACCAAGTGCTTGGGTGGCTCTTGGGCTTA 2640  
|||||  
Db 2581 CTGCACCTCTGAATCTTTACAGATCGTCTTACCAAGTGCTTGGGTGGCTCTTGGGCTTA 2640  
QY 2641 TTGAGATCTTCTTACGAATCATGCCCTCTTGTATGGTATGGTGGCGGTCTGAAT 2700  
|||||  
Db 2641 TTGAGATCTTCTTACGAATCATGCCCTCTTGTATGGTATGGTGGCGGTCTGAAT 2700  
QY 2701 TTTTGAAGATTTTCTACATCACTCCATCGTATCTTGGACATCTATCCCTCT 2760  
|||||  
Db 2701 TTTTGAAGATTTTCTACATCACTCCATCGTATCTTGGACATCTATCCCTCT 2760  
QY 2761 TGGCTTACTGTACATTTGCCCTGCCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGC 2820  
|||||  
Db 2761 TGGCTTACTGTACATTTGCCCTGCCATCTGTTTATTGACAGGAAATTTATCACTCCAGAGC 2820  
QY 2821 TGAATATGTTGCGCAGCCTGTGTTCACTGTCATTTTATCTGCAATTTTGTCTACGAGA 2880  
|||||  
Db 2821 TGAATATGTTGCGCAGCCTGTGTTCACTGTCATTTTATCTGCAATTTTGTCTACGAGA 2880  
QY 2881 TCCTAGAAATGAGATGAGTGTGTGGAATTGATGACTGTGTGGAAGATGAGCAGTTCT 2940  
|||||  
Db 2881 TCCTAGAAATGAGATGAGTGTGTGGAATTGATGACTGTGTGGAAGATGAGCAGTTCT 2940  
QY 2941 GGGTCATTGAGAGTGTGTCTCAACCTCTTGTGTCTTCCAGGACTTCTCAAGTCA 3000  
|||||  
Db 2941 GGGTCATTGAGAGTGTGTCTCAACCTCTTGTGTCTTCCAGGACTTCTCAAGTCA 3000  
QY 3001 TAGCTGTGTGATACAAAGCTTCAACCTGATCATCAAAAGGTGAGATGATGAGAGTTCT 3060  
|||||  
Db 3001 TAGCTGTGTGATACAAAGCTTCAACCTGATCATCAAAAGGTGAGATGATGAGAGTTCT 3060  
QY 3061 CAGAGCTATATACATTCAATGAGTACTTATTGATACCTTACCACCTTGTCTTAT 3120  
|||||  
Db 3061 CAGAGCTATATACATTCAATGAGTACTTATTGATACCTTACCACCTTGTCTTAT 3120  
QY 3121 TGAATTCATTGTTGTGTGCGTGGCGTTTCAATGCGATCAATAACGGATATGATCAT 3180  
|||||  
Db 3121 TGAATTCATTGTTGTGTGCGTGGCGTTTCAATGCGATCAATAACGGATATGATCAT 3180  
QY 3181 GGGGCCCCCTCTTGGGAAGCTATTCTTGCATTGTTGGGTGATGTCCATTATCCCT 3240  
|||||  
Db 3181 GGGGCCCCCTCTTGGGAAGCTATTCTTGCATTGTTGGGTGATGTCCATTATCCCT 3240  
QY 3241 TTCTCAAGGTTTGTGGAAGCAAAACAGACACCAACGATGTCTGTTGTTCA 3300  
|||||  
Db 3241 TTCTCAAGGTTTGTGGAAGCAAAACAGACACCAACGATGTCTGTTGTTCA 3300  
QY 3301 TTCTGTGCTTCAATCTTCTGCTCTTGGGTTGCGATGATCCTTCTGCGAAG 3360  
|||||  
Db 3301 TTCTGTGCTTCAATCTTCTGCTCTTGGGTTGCGATGATCCTTCTGCGAAG 3360  
QY 3361 ATGATGCTCGCTTCTTGAAGAGTGTGTTGATGCACTAGGATGTCAGTGATCAG 3420  
|||||  
Db 3361 ATGATGCTCGCTTCTTGAAGAGTGTGTTGATGCACTAGGATGTCAGTGATCAG 3420  
QY 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTTCTGTGTTTGTCCCCCATATTCA 3480  
|||||

Db 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGTGTTTGTCCCCCATATTCA 3480  
QY 3481 CTGTAGATAGAGACATGAAATGTCCCAAGTTCTTTTGATCCATGTTGAACCTTA 3540  
|||||  
Db 3481 CTGTAGATAGAGACATGAAATGTCCCAAGTTCTTTTGATCCATGTTGAACCTTA 3540  
QY 3541 TATCTGAGATATACTGGGGGAAATGAGGCTGCGGCAATCCTTGTGACGTTGGGCG 3600  
|||||  
Db 3541 TATCTGAGATATACTGGGGGAAATGAGGCTGCGGCAATCCTTGTGACGTTGGGCG 3600  
QY 3601 TGAATACAGCATATGCAAGTGTGTTGATGTGACCATCTTATTAATCTTGTGCAATA 3660  
|||||  
Db 3601 TGAATACAGCATATGCAAGTGTGTTGATGTGACCATCTTATTAATCTTGTGCAATA 3660  
QY 3661 TAGATGGGCTGAGCCGAAACAGCAAGTATTTGATTTCTGCACTGCTCCCGTGAACA 3720  
|||||  
Db 3661 TAGATGGGCTGAGCCGAAACAGCAAGTATTTGATTTCTGCACTGCTCCCGTGAACA 3720  
QY 3721 TGGTTCATTAAGGACGCAAGATGCATCTGCCAGTGAACAGAGCAACCTGACATT 3780  
|||||  
Db 3721 TGGTTCATTAAGGACGCAAGATGCATCTGCCAGTGAACAGAGCAACCTGACATT 3780  
QY 3781 ATTTATGTATGCTGTTCATTGGAGGGCTGTTCATTACATGTTCTGTATAGTAA 3840  
|||||  
Db 3781 ATTTATGTATGCTGTTCATTGGAGGGCTGTTCATTACATGTTCTGTATAGTAA 3840  
QY 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAGTATGTAATGCGCTGTTTGT 3900  
|||||  
Db 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAGTATGTAATGCGCTGTTTGT 3900  
QY 3901 TGTGTACTGTAATCATCTGAGTTGTTTGTGAAA 3936  
|||||  
Db 3901 TGTGTACTGTAATCATCTGAGTTGTTTGTGAAA 3936  
|||||  
RESULT 2  
US-10-209-059-49  
; Sequence 49, Application US/10209059  
; Publication No. US20030163838A1  
; GENERAL INFORMATION:  
; APPLICANT: Dnuga, Kanwarpal S.  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 0864R2  
; CURRENT APPLICATION NUMBER: US/10/209,059  
; PRIOR FILING DATE: 2002-07-31  
; PRIOR APPLICATION NUMBER: 60/096,822  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 09/371,383  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR APPLICATION NUMBER: 09/550,483  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49  
; LENGTH: 3969  
; TYPE: DNA  
; ORGANISM: Zea mays  
; US-10-209-059-49  
Query Match 100.0%; Score 3936; DB 13; Length 3969;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCTCCCTCGTGGTGGCGCGTGGCGCGGCTTCCGTGAGAAACCACTGGGGG 60  
|||||  
Db 1 CTCTCCCTCGTGGTGGCGCGTGGCGCGGCTTCCGTGAGAAACCACTGGGGG 60  
QY 61 ATGAGATCTGCTGCTAGAGTGAAGAGAGTACGTCAGTATCTCTGCTTCTGCGG 120  
|||||  
Db 61 ATGAGATCTGCTGCTAGAGTGAAGAGAGTACGTCAGTATCTCTGCTTCTGCGG 120  
|||||

QY 121 GCCGAAGTGGAGGGGAGGAAAGCGATGAGAGCGGAGCGCGGCGTGTGGCCGGCTCCACA 180  
|  
Db 121 GCCGAAGTGGAGGGGAGGAAAGCGATGAGAGCGGAGCGCGGCGTGTGGCCGGCTCCACA 180  
QY 181 ACCGCAACGAGCTCGTCGTATCCGCGCGACGCGGATCCCGGCGAAAGCCGCGCGG 240  
|  
Db 181 ACCGCAACGAGCTCGTCGTATCCGCGCGACGCGGATCCCGGCGAAAGCCGCGCGG 240  
QY 241 AGCAGAACGGGCAAGTGTGCCAGATTTGGCGCGACGACGTGGCCTTGCCCCGCGGGG 300  
|  
Db 241 AGCAGAACGGGCAAGTGTGCCAGATTTGGCGCGACGACGTGGCCTTGCCCCGCGGGG 300  
QY 301 ACCCCTTCGTGGCGGTGCAACGAGTGCCTTCCCCGTCTGCGGGGACTGTCTACGAATACG 360  
|  
Db 301 ACCCCTTCGTGGCGGTGCAACGAGTGCCTTCCCCGTCTGCGGGGACTGTCTACGAATACG 360  
QY 361 AGCGCCGGGAGGGCAAGCAAGAACTGCCCCAGTGCAAGACTCGATACAAAGCGCTCAAGG 420  
|  
Db 361 AGCGCCGGGAGGGCAAGCAAGAACTGCCCCAGTGCAAGACTCGATACAAAGCGCTCAAGG 420  
QY 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGGACGCGCGTGCATGACTGACCAACGAGT 480  
|  
Db 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGGACGCGCGTGCATGACTGACCAACGAGT 480  
QY 481 TCAACTGGGACGGCCATGACTCGAGTCTGTGCGCGAGTCCATGCTCTACGCCACATGA 540  
|  
Db 481 TCAACTGGGACGGCCATGACTCGAGTCTGTGCGCGAGTCCATGCTCTACGCCACATGA 540  
QY 541 GCTACGGCCGTGGAGGTGACCCCTAATGGCGCGCCACAGCTTTCAGCTCAACCCCAATG 600  
|  
Db 541 GCTACGGCCGTGGAGGTGACCCCTAATGGCGCGCCACAGCTTTCAGCTCAACCCCAATG 600  
QY 601 TTCCACTCTCTACCAACGGGCAATGGTGAATGACATCCCAACCGGACAGCAGCAGCGCTGG 660  
|  
Db 601 TTCCACTCTCTACCAACGGGCAATGGTGAATGACATCCCAACCGGACAGCAGCAGCGCTGG 660  
QY 661 TGCCTTCTTCATGGGTGGTGGGGAAGGATACATCCCTTCCTTATGCGGATCCCA 720  
|  
Db 661 TGCCTTCTTCATGGGTGGTGGGGAAGGATACATCCCTTCCTTATGCGGATCCCA 720  
QY 721 GCTTACCTGTGCAACCCAGGCTATGAGACCCATCCAAGATCTTGCTGATATGGGTATG 780  
|  
Db 721 GCTTACCTGTGCAACCCAGGCTATGAGACCCATCCAAGATCTTGCTGATATGGGTATG 780  
QY 781 GTAGTGTGCTTGAAGGAACGATGGAGAATTGGAAGCAGAGACAAGAGATGCACC 840  
|  
Db 781 GTAGTGTGCTTGAAGGAACGATGGAGAATTGGAAGCAGAGACAAGAGATGCACC 840  
QY 841 AGACGGGGAATGATGTTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATG 900  
|  
Db 841 AGACGGGGAATGATGTTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 AAGCAAGACAACAACGTGCCAGAAATTCACCTTCATCAAGCAGATTAATCCATATA 960  
|  
Db 901 AAGCAAGACAACAACGTGCCAGAAATTCACCTTCATCAAGCAGATTAATCCATATA 960  
QY 961 GGATGATTAATCATTAATTCGGCTTGTGGTTTGGGGTCTTCTTCACTACCGAGTATGC 1020  
|  
Db 961 GGATGATTAATCATTAATTCGGCTTGTGGTTTGGGGTCTTCTTCACTACCGAGTATGC 1020  
QY 1021 ATCCGGTGAATGATGATTTGCTTGTGGCTATATCTGTTATCTGTGAATCTGTTTG 1080  
|  
Db 1021 ATCCGGTGAATGATGATTTGCTTGTGGCTATATCTGTTATCTGTGAATCTGTTTG 1080  
QY 1081 CCATGTCTTGATTTCTTGATCAATTCGCAAAAGTGTTCCCTATGAGAGAGACTTACC 1140  
|  
Db 1081 CCATGTCTTGATTTCTTGATCAATTCGCAAAAGTGTTCCCTATGAGAGAGACTTACC 1140  
QY 1141 TAGACCGGCTGTCACTGAGGTTGCAACAAGAAAGCCAGCCATCTCAACTTGCTCCAATTG 1200  
|  
Db 1141 TAGACCGGCTGTCACTGAGGTTGCAACAAGAAAGCCAGCCATCTCAACTTGCTCCAATTG 1200

QY 1201 ATTCTTTGTCACTACGGTGTATCCCTTAAAGAACTCCTTTGGTCAACAATACTG 1260  
|  
Db 1201 ATTCTTTGTCACTACGGTGTATCCCTTAAAGAACTCCTTTGGTCAACAATACTG 1260  
QY 1261 TTCTATCTATCCTTTCGGTGGATATATCCTGTGTGATAAGGTTTCTGTATGTTCTGATG 1320  
|  
Db 1261 TTCTATCTATCCTTTCGGTGGATATATCCTGTGTGATAAGGTTTCTGTATGTTCTGATG 1320  
QY 1321 ATGTGCTGCAATGCTTAACGTTTGAAGCATTAATGAAACATCTGAATTTGCAAGAAAT 1380  
|  
Db 1321 ATGTGCTGCAATGCTTAACGTTTGAAGCATTAATGAAACATCTGAATTTGCAAGAAAT 1380  
QY 1381 GGGTTCCTTCTGCAACGGTACAAATATTGAACCTGCGCTCCAGAGTGTACTTCCAAC 1440  
|  
Db 1381 GGGTTCCTTCTGCAACGGTACAAATATTGAACCTGCGCTCCAGAGTGTACTTCCAAC 1440  
QY 1441 AGAAGATAGACTACTTGAAGACAAGGTGCGAGCAAACTTGTGAAGAGAGAGACAA 1500  
|  
Db 1441 AGAAGATAGACTACTTGAAGACAAGGTGCGAGCAAACTTGTGAAGAGAGAGACAA 1500  
QY 1501 TGAAGAGAGATATGAGAAATTCAAGGTGAGAAATCAATGCTTAGTTGCCAAAGCCAGA 1560  
|  
Db 1501 TGAAGAGAGATATGAGAAATTCAAGGTGAGAAATCAATGCTTAGTTGCCAAAGCCAGA 1560  
QY 1561 AAGTTCCTGAAGAGATGGAACAATGCAAGATGGAACCCCTGGCCTGGAACAATGTTTC 1620  
|  
Db 1561 AAGTTCCTGAAGAGATGGAACAATGCAAGATGGAACCCCTGGCCTGGAACAATGTTTC 1620  
QY 1621 GTGATCATCTCGGAATGATTCAAGTCTTCTTGCCAAAGCGGAGGCTTGACTGTGAGG 1680  
|  
Db 1621 GTGATCATCTCGGAATGATTCAAGTCTTCTTGCCAAAGCGGAGGCTTGACTGTGAGG 1680  
QY 1681 GAAATGAACCTGCCAGATGGTTTATGTTTCTAGAGAGAAACGACAGGCTATTAACCATC 1740  
|  
Db 1681 GAAATGAACCTGCCAGATGGTTTATGTTTCTAGAGAGAAACGACAGGCTATTAACCATC 1740  
QY 1741 ATAAGAAAAGCTGTCATGAATGATGATGTCGCGAGTCTGCTGTACTTAACAATGCTC 1800  
|  
Db 1741 ATAAGAAAAGCTGTCATGAATGATGATGTCGCGAGTCTGCTGTACTTAACAATGCTC 1800  
QY 1801 CATATTTGTTAACTTGAATTTGTGATCACTACTACATCAACAAGCAAGGCTATTAAGGAAG 1860  
|  
Db 1801 CATATTTGTTAACTTGAATTTGTGATCACTACTACATCAACAAGCAAGGCTATTAAGGAAG 1860  
QY 1861 CAATGTGTTTATGATGACCCCTTACTAGGAAGAAAGGTTGCTATGTACAGTTCCTC 1920  
|  
Db 1861 CAATGTGTTTATGATGACCCCTTACTAGGAAGAAAGGTTGCTATGTACAGTTCCTC 1920  
QY 1921 AAAGATTTGATGGATGATTCGCCATGACCGATATGCTAAACCGAATGTTGTTTTTG 1980  
|  
Db 1921 AAAGATTTGATGGATGATTCGCCATGACCGATATGCTAAACCGAATGTTGTTTTTG 1980  
QY 1981 ATATCAACATGAAGGTTTGGATGTTTCAAGGTCCAATTTATGTTGGTACTGATGTG 2040  
|  
Db 1981 ATATCAACATGAAGGTTTGGATGTTTCAAGGTCCAATTTATGTTGGTACTGATGTG 2040  
QY 2041 TATTTAGAAGGCAAGCATTAATGTTATGATGATGATGATGATGATGATGATGATGATG 2100  
|  
Db 2041 TATTTAGAAGGCAAGCATTAATGTTATGATGATGATGATGATGATGATGATGATGATG 2100  
QY 2101 GGACTTGCAACTGCTGGGCCAAGTGTGCTTTTGTGCTGTGCTGCTTGGCAATAGGAAGC 2160  
|  
Db 2101 GGACTTGCAACTGCTGGGCCAAGTGTGCTTTTGTGCTGTGCTGCTTGGCAATAGGAAGC 2160  
QY 2161 AAAAGAAAGACTACCAAAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
|  
Db 2161 AAAAGAAAGACTACCAAAACCCAAACAGAGAAAGAAAGTTATTTTCAAGAAAGAG 2220  
QY 2221 AGAACAATCCCTGCAATATGCTCTTGTGAATTTGACGAAGCTGCTCCAGAGCTGAGA 2280  
|  
Db 2221 AGAACAATCCCTGCAATATGCTCTTGTGAATTTGACGAAGCTGCTCCAGAGCTGAGA 2280  
QY 2281 ATGAAAAAGCCGATTTGTAAATCAACAAAAATTAAGAAAAAATTTGGCCAATCTTCTG 2340

```
Db 2281 ATGAAAGCGCGTATGTGAATCAACAAAATTAGAAAAGAAATTGGCCAACTCTCTG 2340
QY 2341 TTTTGTACATCCACACTCTCGAGATGSGAACCTTGAAGAGTGCAGTCTCTG 2400
Db 2341 TTTTGTACATCCACACTCTCGAGATGSGAACCTTGAAGAGTGCAGTCTCTG 2400
QY 2401 CTCTTTGAAAAGCTATACATGTCTATTAGTGTGTTATGAAGACAGACAGCTGGG 2460
Db 2401 CTCTTTGAAAAGCTATACATGTCTATTAGTGTGTTATGAAGACAGACAGCTGGG 2460
QY 2461 GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAGATATTCTAAGTGTCAAGA 2520
Db 2461 GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAGATATTCTAAGTGTCAAGA 2520
QY 2521 TGCATTGTGATGTTGGCGGCTCAATTTACTGCATACCTAAACGGGTGCATTCAAGTT 2580
Db 2521 TGCATTGTGATGTTGGCGGCTCAATTTACTGCATACCTAAACGGGTGCATTCAAGTT 2580
QY 2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAAGTCTCGGTGGGCTCTGGGCTCTA 2640
Db 2581 CTGCACCTCTGAATCTTTCAGATCGTCTTCACCAAGTCTCGGTGGGCTCTGGGCTCTA 2640
QY 2641 TTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGTTGGCGGTCTGAAT 2700
Db 2641 TTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGTTGGCGGTCTGAAT 2700
QY 2701 TTTTGAAGATTTTCCCTACATCAACTCCATCGTGTATCCTTGACATCTATCCCTCT 2760
Db 2701 TTTTGAAGATTTTCCCTACATCAACTCCATCGTGTATCCTTGACATCTATCCCTCT 2760
QY 2761 TGGCTTACTGTACATTTGCCCATCTGTTATTGACAGGAAATTATCACTCCAGAGC 2820
Db 2761 TGGCTTACTGTACATTTGCCCATCTGTTATTGACAGGAAATTATCACTCCAGAGC 2820
QY 2821 TGAATATATGTTGCCAGCCTGTGTTTCATGTCACTTTTATCTGCATTTTGTCTAGAGCA 2880
Db 2821 TGAATATATGTTGCCAGCCTGTGTTTCATGTCACTTTTATCTGCATTTTGTCTAGAGCA 2880
QY 2881 TCCTAGAAATGAGATGAGTGTGTTGAATTGATGACTGSGAGAGATGAGCACTTCT 2940
Db 2881 TCCTAGAAATGAGATGAGTGTGTTGAATTGATGACTGSGAGAGATGAGCACTTCT 2940
QY 2941 GGGTCATTGAGGTGTGCTTCAACCTCTTGTGCTGTCCAGGACTTCTCAAGTCA 3000
Db 2941 GGGTCATTGAGGTGTGCTTCAACCTCTTGTGCTGTCCAGGACTTCTCAAGTCA 3000
QY 3001 TAGCTGTGTGATACAGCTTCAACCGTGACATCAAGGTTGAGATGATGAGAGTCT 3060
Db 3001 TAGCTGTGTGATACAGCTTCAACCGTGACATCAAGGTTGAGATGATGAGAGTCT 3060
QY 3061 CAGAGCTATATACATTCAAAATGGACTACCTTATTGATACCTCCACACCTTGTCTAT 3120
Db 3061 CAGAGCTATATACATTCAAAATGGACTACCTTATTGATACCTCCACACCTTGTCTAT 3120
QY 3121 TGAACCTTATTGTTGTGCTGCTGGCGTTTCAAAATGCGATCAATAACGAGATGAGTCA 3180
Db 3121 TGAACCTTATTGTTGTGCTGCTGGCGTTTCAAAATGCGATCAATAACGAGATGAGTCA 3180
QY 3181 GGGGCCCCCTCTTTGGAGAGCTATTCTTTGCAATTTGGGATTTGCCATCTTATCCCT 3240
Db 3181 GGGGCCCCCTCTTTGGAGAGCTATTCTTTGCAATTTGGGATTTGCCATCTTATCCCT 3240
QY 3241 TTCTCAAAGTTTGTGGAAGCAAAACAGACACCAAGATTGTCACTGCTGCTCCA 3300
Db 3241 TTCTCAAAGTTTGTGGAAGCAAAACAGACACCAAGATTGTCACTGCTGCTCCA 3300
QY 3301 TTCTGCTGCTTCAATCTTCTGCTCTTGGGTTGGATTTGATCCTTTCCTTCCGAAAG 3360
Db 3301 TTCTGCTGCTTCAATCTTCTGCTCTTGGGTTGGATTTGATCCTTTCCTTCCGAAAG 3360
QY 3361 ATGATGTCGCTTCTTGAGAGATGTGTTTGATTCGAATGATGTCACTGATCAG 3420
Db 3361 ATGATGTCGCTTCTTGAGAGATGTGTTTGATTCGAATGATGTCACTGATCAG 3420
```

```
Db 3361 ATGATGTCGCTTCTTGAGAGATGTGTTTGATTCGAATGATGTCACTGATCAG 3420
QY 3421 CTCCCCCAATCTGATATGCTTGAAGTATAATTCTGTGTTGTGCCCATATTCACTGT 3480
Db 3421 CTCCCCCAATCTGATATGCTTGAAGTATAATTCTGTGTTGTGCCCATATTCACTGT 3480
QY 3481 CTGTAGATAAGACATGAATGTCCCAAGTCTTTTGATCCATGAGTAACTTAA 3540
Db 3481 CTGTAGATAAGACATGAATGTCCCAAGTCTTTTGATCCATGAGTAACTTAA 3540
QY 3541 TATCTGAGATATACTGGGGGAAAATGAGGCTGCGCAATCCTTGTGAGTTGGCCG 3600
Db 3541 TATCTGAGATATACTGGGGGAAAATGAGGCTGCGCAATCCTTGTGAGTTGGCCG 3600
QY 3601 TGAATACAGCATATGCAAGTGTGATTGTCAGACATTCTTATTACTTGTGCAATA 3660
Db 3601 TGAATACAGCATATGCAAGTGTGATTGTCAGACATTCTTATTACTTGTGCAATA 3660
QY 3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTGATTCTGCACCTGCTCCGTTCAAACT 3720
Db 3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTGATTCTGCACCTGCTCCGTTCAAACT 3720
QY 3721 TGGTCTCAATAGGCAAGGCAAGATGCACTGCCAGTGGAACAGACAACTGCACATT 3780
Db 3721 TGGTCTCAATAGGCAAGGCAAGATGCACTGCCAGTGGAACAGACAACTGCACATT 3780
QY 3781 ATTTATGTATGCTGCTTTCATTGAGGGCTGTTCATTACATGTTGCTATACTAGAAA 3840
Db 3781 ATTTATGTATGCTGCTTTCATTGAGGGCTGTTCATTACATGTTGCTATACTAGAAA 3840
QY 3841 AACAGATATATAGCATTAATCTATAGTTAATTAAGTATGTAATGCGCTGTTTGT 3900
Db 3841 AACAGATATATAGCATTAATCTATAGTTAATTAAGTATGTAATGCGCTGTTTGT 3900
QY 3901 TGTGTACTGTAATCATCTGAGTTGTTGTGAAA 3936
Db 3901 TGTGTACTGTAATCATCTGAGTTGTTGTGAAA 3936
```

```
RESULT 3
US-10-160-719-17
; Sequence 17, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR APPLICATION NUMBER: US/09/371,363A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 3969
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144) ... (3401)
US-10-160-719-17
```

```
Query Match 100.0%; Score 3936; DB 13; Length 3969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CTCTCCCTGTCGGTGGCGGCTGCGCGGCTCGGCGTTCGGTGAAGAACTCGGGGG 60
```



|||||  
Db 1 CTTCTCCCTGTCGGTGC GGCGCTCGGCTCGGCTCGGTGAGAAACCACTCGGGG 60  
QY 61 ATGAGATCTGCTAGTAGAGAGCTACGGTCAGTATCCTCTGCTTCTGTCGGC 120  
Db 61 ATGAGATCTGCTAGTAGAGAGAGCTACGGTCAGTATCCTCTGCTTCTGTCGGC 120  
QY 121 GCGGAAGTGGAGGGAGGAAGCGATGGAGCGAGCGCGGCTGTGTGCGCTCCACA 180  
Db 121 GCGGAAGTGGAGGGAGGAAGCGATGGAGCGAGCGCGGCTGTGTGCGCTCCACA 180  
QY 181 ACCGCAACGACTCGTCATCCGCGCAGCGCGATCCCGGCGGAAGCGCGCGG 240  
Db 181 ACCGCAACGACTCGTCATCCGCGCAGCGCGATCCCGGCGGAAGCGCGCGG 240  
QY 241 AGCAGAACGGGAGGTGTGCCAGATTGGCGCGACGACGTGGCTTGCCCCGCGGG 300  
Db 241 AGCAGAACGGGAGGTGTGCCAGATTGGCGCGACGACGTGGCTTGCCCCGCGGG 300  
QY 301 ACCCCTTCTGCGGTGCAACGAGTGCCTTCCCGTCTGCCGGGACTGCTACGAATACG 360  
Db 301 ACCCCTTCTGCGGTGCAACGAGTGCCTTCCCGTCTGCCGGGACTGCTACGAATACG 360  
QY 361 AGCGCCGGGAGGCGACGCAAGACTGCCCCCAGTCAAGACTCGATACAGCGCCTCAAG 420  
Db 361 AGCGCCGGGAGGCGACGCAAGACTGCCCCCAGTCAAGACTCGATACAGCGCCTCAAG 420  
QY 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGAGCGGCGTGCATGCTGGAACAAGT 480  
Db 421 GCTGCCAACGTGTGACCGGTGACGAGAGAGAGCGGCGTGCATGCTGGAACAAGT 480  
QY 481 TCAACTGGAGCGGCCATGACTCGCAGTCTGTGCGCGAGTCCATGCTCTACGGCCACATGA 540  
Db 481 TCAACTGGAGCGGCCATGACTCGCAGTCTGTGCGCGAGTCCATGCTCTACGGCCACATGA 540  
QY 541 GCTACGCGGTGAGGTGACCCCTAATGCGCGCGCACAAAGCTTTCAGCTCAACCCCAATG 600  
Db 541 GCTACGCGGTGAGGTGACCCCTAATGCGCGCGCACAAAGCTTTCAGCTCAACCCCAATG 600  
QY 601 TTCCACTCTCAACGCGGCAATGTGTGATGACATCCCAACGAGCAGCAGCGCTGG 660  
Db 601 TTCCACTCTCTCAACGCGGCAATGTGTGATGACATCCCAACGAGCAGCAGCGCTGG 660  
QY 661 TGCCTTCTTTCATGGGTGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCCA 720  
Db 661 TGCCTTCTTTCATGGGTGTGGGGGAAAGAGATACATCCCTTCTTATGCGGATCCCA 720  
QY 721 GCTTACCTGTCAACCCAGGTCTATGACCCATCCAAGATCTGTGCTCATATGGGTATG 780  
Db 721 GCTTACCTGTCAACCCAGGTCTATGACCCATCCAAGATCTGTGCTCATATGGGTATG 780  
QY 781 GTAGTGTGCTTGAAGGAACGGATGAGAAATTGAAGCAGAGCAAGAGAGATGCACC 840  
Db 781 GTAGTGTGCTTGAAGGAACGGATGAGAAATTGAAGCAGAGCAAGAGAGATGCACC 840  
QY 841 AGACGGGAGATGATGCTGTGTGATGATGCTGACATGCTATACCACTAATGATG 900  
Db 841 AGACGGGAGATGATGCTGTGTGATGATGCTGACATGCTATACCACTAATGATG 900  
QY 901 AAGCAAGACAACACTGTCCAGGAAATTCACCTTCCATCAAGCCAGATTAATCCATATA 960  
Db 901 AAGCAAGACAACACTGTCCAGGAAATTCACCTTCCATCAAGCCAGATTAATCCATATA 960  
QY 961 GGATGATTATCATTAATCGGCTTGTGTTTGGGTTCTTCTTCCACTACGAGTGATGC 1020  
Db 961 GGATGATTATCATTAATCGGCTTGTGTTTGGGTTCTTCTTCCACTACGAGTGATGC 1020  
QY 1021 ATCCGGTGAATGATGCAATTGCTTGTGCTCATATCTGTTATCTGTAATCTGTTTG 1080  
Db 1021 ATCCGGTGAATGATGCAATTGCTTGTGCTCATATCTGTTATCTGTAATCTGTTTG 1080  
QY 1081 CCATGTCTTGATCTTGATCAATTCCCAAGTGTCCCTATTGAGAGAGACTTACC 1140  
Db 1081 CCATGTCTTGATCTTGATCAATTCCCAAGTGTGTCCCTATTGAGAGAGACTTACC 1140

Db 1081 CCATGTCTTGATCTTGATCAATTCCCAAGTGTGTCCCTATTGAGAGAGACTTACC 1140  
QY 1141 TAGACCGGCTGTACTGAGGTTGCAACAAGGAGGCGACCATCTCAACTTGTCTCAATTG 1200  
Db 1141 TAGACCGGCTGTACTGAGGTTGCAACAAGGAGGCGACCATCTCAACTTGTCTCAATTG 1200  
QY 1201 ATTCTTTGTCAGTACGTTGATCCCTTAAGGAACCTCTTGTGTACAACAATACTG 1260  
Db 1201 ATTCTTTGTCAGTACGTTGATCCCTTAAGGAACCTCTTGTGTACAACAATACTG 1260  
QY 1261 TTCTATCTATCCTTTCGGTGATTAATCCTGTGTAAGGTTTCTGTATGTTCTGATG 1320  
Db 1261 TTCTATCTATCCTTTCGGTGATTAATCCTGTGTGTAAGGTTTCTGTATGTTCTGATG 1320  
QY 1321 ATGCTGCTCAATGCTTAACGTTTGAAGCATTAATGAAACATCTGAATTTGCAAGAAT 1380  
Db 1321 ATGCTGCTCAATGCTTAACGTTTGAAGCATTAATGAAACATCTGAATTTGCAAGAAT 1380  
QY 1381 GGGTTCCTTCTGCAAAACGTTCAATATTTGAACCTTCGCGCTCCAGAGTGTACTTCCAA 1440  
Db 1381 GGGTTCCTTCTGCAAAACGTTCAATATTTGAACCTTCGCGCTCCAGAGTGTACTTCCAA 1440  
QY 1441 AGAAGATAGACTACTTGAAAGCAAGGTGGCAGCAAACTTGTTAGGAGAGAGCAAA 1500  
Db 1441 AGAAGATAGACTACTTGAAAGCAAGGTGGCAGCAAACTTGTTAGGAGAGAGCAAA 1500  
QY 1501 TGAAGAGAGATATGAGGAATTCAAGGTGAGATCAATGCTTAGTTGCCAAAGCCAGA 1560  
Db 1501 TGAAGAGAGATATGAGGAATTCAAGGTGAGATCAATGCTTAGTTGCCAAAGCCAGA 1560  
QY 1561 AAGTTCCTGAAGAAGATGCAATGCAAGATGAAACCCCTGCGCTGGAACAATGTTT 1620  
Db 1561 AAGTTCCTGAAGAAGATGCAATGCAAGATGAAACCCCTGCGCTGGAACAATGTTT 1620  
QY 1621 GTGATCATCCTGGAATGATTCAGGTCTTCTTGGCCAAAGCGGAGGCTTGACTGTGAGG 1680  
Db 1621 GTGATCATCCTGGAATGATTCAGGTCTTCTTGGCCAAAGCGGAGGCTTGACTGTGAGG 1680  
QY 1681 GAAATGAACCTGCCAGATTTGTTATGTTCTTAGAGAGAAACGACGAGCTATTAACATC 1740  
Db 1681 GAAATGAACCTGCCAGATTTGTTATGTTCTTAGAGAGAAACGACGAGCTATTAACATC 1740  
QY 1741 ATTAAGAAAGCTGTGCTATGATGATGCTTCCAGTCTGCTGTACTTAACAAATGCTC 1800  
Db 1741 ATTAAGAAAGCTGTGCTATGATGATGCTTCCAGTCTGCTGTACTTAACAAATGCTC 1800  
QY 1801 CATATTTGTTAACTTGATTTGATCACTACATCAACAACAGAGCTATAAAGAG 1860  
Db 1801 CATATTTGTTAACTTGATTTGATCACTACATCAACAACAGAGCTATAAAGAG 1860  
QY 1861 CAATGTGTTTATGATGAGACCTTTACTAGGAAAGAGGTTGCTATGTACAGTTCCCTC 1920  
Db 1861 CAATGTGTTTATGATGAGACCTTTACTAGGAAAGAGGTTGCTATGTACAGTTCCCTC 1920  
QY 1921 AAAGATTTGATGGGATGATGCGCATGACCGATATGCTAACCGGAATGTGTCTTTTGG 1980  
Db 1921 AAAGATTTGATGGGATGATGCGCATGACCGATATGCTAACCGGAATGTGTCTTTTGG 1980  
QY 1981 ATATCAACATGAAGGTTTGAATGATTTAGGGTCCAAATTTATGTTGTTACTGATGTG 2040  
Db 1981 ATATCAACATGAAGGTTTGAATGATTTAGGGTCCAAATTTATGTTGTTACTGATGTG 2040  
QY 2041 TATTTAGAGGCAAGCATTAATGTTATGATGCCCCCAAAACAAGAGCCACCATCA 2100  
Db 2041 TATTTAGAGGCAAGCATTAATGTTATGATGCCCCCAAAACAAGAGCCACCATCA 2100  
QY 2101 GGAATTGCAACTGTGCGCCCAAGTGTCTTTTGTGCTGTCTTTGGCAATAGGAAGC 2160  
Db 2101 GGAATTGCAACTGTGCGCCCAAGTGTCTTTTGTGCTGTCTTTGGCAATAGGAAGC 2160  
QY 2161 AAAAGAGACTACCAAAACCAAAACAGAGAGAAAAAGTTATTTTCAAGAAAGAG 2220  
Db 2161 AAAAGAGACTACCAAAACCAAAACAGAGAGAAAAAGTTATTTTCAAGAAAGAG 2220

OY 2221 AGAACCAATCCCTGCATATGCTCTTGGAATTGACGAAGCTCTCCAGAGCTGAGA 2280  
|||||  
Db 2221 AGAACCAATCCCTGCATATGCTCTTGGAATTGACGAAGCTCTCCAGAGCTGAGA 2280  
OY 2281 ATGAAAAGCCGGTATTGTAATCAACAAAAATTAGAAAAGAAATTTGGCCAACTCTTG 2340  
|||||  
Db 2281 ATGAAAAGCCGGTATTGTAATCAACAAAAATTAGAAAAGAAATTTGGCCAACTCTTG 2340  
OY 2341 TTTTGTATACATCCACACTCTCGAATGCTGGAACCTTGAAGATGCAAGTCTGCTT 2400  
|||||  
Db 2341 TTTTGTATACATCCACACTCTCGAATGCTGGAACCTTGAAGATGCAAGTCTGCTT 2400  
OY 2401 CTCTTTGAAAGAAGCTATACATGCTAGTGTGTTATGAAAGACAAGACAGACTGGG 2460  
|||||  
Db 2401 CTCTTTGAAAGAAGCTATACATGCTAGTGTGTTATGAAAGACAAGACAGACTGGG 2460  
OY 2461 GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAGATATTCTAATCTGTTCAAGA 2520  
|||||  
Db 2461 GAAAAGAGATTGGCTGATCTATGATCAGTTACAGAGATATTCTAATCTGTTCAAGA 2520  
OY 2521 TGCAATGTATGCTTGGCGGTCAATTTACTGCATACCTAAACGGGTGCATTCAAGGTT 2580  
|||||  
Db 2521 TGCAATGTATGCTTGGCGGTCAATTTACTGCATACCTAAACGGGTGCATTCAAGGTT 2580  
OY 2581 CTGACCTCTGAATCTTTCAGATCGTCTTACACAGGTGCTTGGGGCTCTTGAGCTTA 2640  
|||||  
Db 2581 CTGACCTCTGAATCTTTCAGATCGTCTTACACAGGTGCTTGGGGCTCTTGAGCTTA 2640  
OY 2641 TTGAGATCTTCTCAGCAATCATTGCCCTTTGGTATGGGTATGGCGGTCTGAAT 2700  
|||||  
Db 2641 TTGAGATCTTCTCAGCAATCATTGCCCTTTGGTATGGGTATGGCGGTCTGAAT 2700  
OY 2701 TTTTGGAAAGATTTTCTCTACATCACTCAATCGTGTATCTTGACATCTATTCCTCT 2760  
|||||  
Db 2701 TTTTGGAAAGATTTTCTCTACATCACTCAATCGTGTATCTTGACATCTATTCCTCT 2760  
OY 2761 TGGCTTACTGTACATTTGCCCTGCTGCTTATTTGACAGGAAATTTATCACTCCAGAGC 2820  
|||||  
Db 2761 TGGCTTACTGTACATTTGCCCTGCTGCTTATTTGACAGGAAATTTATCACTCCAGAGC 2820  
OY 2821 TGAATAATGTGCGCAGCCTGTGTTCAATGTCACCTTTTATCTGCATTTTGTCTACGACA 2880  
|||||  
Db 2821 TGAATAATGTGCGCAGCCTGTGTTCAATGTCACCTTTTATCTGCATTTTGTCTACGACA 2880  
OY 2881 TCCTAGAAATGAGATGAGTGTGTTGGAATTGATGACTGTGTGAGGAATGAGCAGTTCT 2940  
|||||  
Db 2881 TCCTAGAAATGAGATGAGTGTGTTGGAATTGATGACTGTGTGAGGAATGAGCAGTTCT 2940  
OY 2941 GGGTCATTGGAAGGTGTGTCTCTCACACCTCTTGTGCTGTGTTCCAGGGAATCTCAAGGTCA 3000  
|||||  
Db 2941 GGGTCATTGGAAGGTGTGTCTCTCACACCTCTTGTGCTGTGTTCCAGGGAATCTCAAGGTCA 3000  
OY 3001 TAGCTGTGTGATACAAAGCTTACCGTGACATCAAAAGGGTGAGATGATGAGAGTTCT 3060  
|||||  
Db 3001 TAGCTGTGTGATACAAAGCTTACCGTGACATCAAAAGGGTGAGATGATGAGAGTTCT 3060  
OY 3061 CAGAGCTATATACATTCAATGAGACTTATTTGATACCTCTACCACTTGCTTCTAT 3120  
|||||  
Db 3061 CAGAGCTATATACATTCAATGAGACTTATTTGATACCTCTACCACTTGCTTCTAT 3120  
OY 3121 TGAATTCATTGGTGTGTGCTGGCTTTCAATGCGATCAATAACGATATGAGTCAAT 3180  
|||||  
Db 3121 TGAATTCATTGGTGTGTGCTGGCTTTCAATGCGATCAATAACGATATGAGTCAAT 3180  
OY 3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCAATTTGGGTGATGTCCATCTTATCCCT 3240  
|||||  
Db 3181 GGGGCCCCCTCTTTGGGAAGCTATTCTTTGCAATTTGGGTGATGTCCATCTTATCCCT 3240  
OY 3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGACACCAACGATGTCTATGCTGCTCA 3300  
|||||  
Db 3241 TTCTCAAAGGTTTGGTTGGAAGGCAAAACAGACACCAACGATGTCTATGCTGCTCA 3300

OY 3301 TTCTGCTGGCTTCAATCTTCTGCTCCTTTGGGTGGAATGATCCTTTCCTTGCGAAG 3360  
|||||  
Db 3301 TTCTGCTGGCTTCAATCTTCTGCTCCTTTGGGTGGAATGATCCTTTCCTTGCGAAG 3360  
OY 3361 ATGATGGTCCGCTTCTTGAGAGTGTGTTGGAATTGCACTAGAGTGCAGTGCATCAG 3420  
|||||  
Db 3361 ATGATGGTCCGCTTCTTGAGAGTGTGTTGGAATTGCACTAGAGTGCAGTGCATCAG 3420  
OY 3421 CTCGCCCAATCTGCATATGCTTGAAGTATATTTCTGCTGTTTGTCCCATATTCAGTGT 3480  
|||||  
Db 3421 CTCGCCCAATCTGCATATGCTTGAAGTATATTTCTGCTGTTTGTCCCATATTCAGTGT 3480  
OY 3481 CTGTAGATTAAGACATGAATGTCCCAAGTTCTTTGATCCATGTGTAACCTACTTAA 3540  
|||||  
Db 3481 CTGTAGATTAAGACATGAATGTCCCAAGTTCTTTGATCCATGTGTAACCTACTTAA 3540  
OY 3541 TATCTGAGATATATCTGGGGAATAATGAGGCTGCGCAATCCTTGTCAGTTGGCCG 3600  
|||||  
Db 3541 TATCTGAGATATATCTGGGGAATAATGAGGCTGCGCAATCCTTGTCAGTTGGCCG 3600  
OY 3601 TGAATACAGCATATGCAAGTGTGATTTGATGTGACAGATCTTATTAATCTTGTGCAATA 3660  
|||||  
Db 3601 TGAATACAGCATATGCAAGTGTGATTTGATGTGACAGATCTTATTAATCTTGTGCAATA 3660  
OY 3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTGATTCTGCACTGCTCCCGTGTACAACT 3720  
|||||  
Db 3661 TAGATGGGCTGAGCCGAACAGCAAGGTATTTGATTCTGCACTGCTCCCGTGTACAACT 3720  
OY 3721 TGGTCTCAATAAAGCAGGCAAGATGATCTGCGAGTGAACAGACAACTGACAT 3780  
|||||  
Db 3721 TGGTCTCAATAAAGCAGGCAAGATGATCTGCGAGTGAACAGACAACTGACAT 3780  
OY 3781 ATTTATGTATGCCCTGTTCAATTGGAAGGCTTGTTCATTAATGTTCTATATAGAAAA 3840  
|||||  
Db 3781 ATTTATGTATGCCCTGTTCAATTGGAAGGCTTGTTCATTAATGTTCTATATAGAAAA 3840  
OY 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAAGTATGTAATGCCCTGTTTGT 3900  
|||||  
Db 3841 AACAGATATTAGCATTAATCTATAGTTAATTAAAGTATGTAATGCCCTGTTTGT 3900  
OY 3901 TGTGTACTGTATCATCTGAGTTGTTGTGAAAA 3936  
|||||  
Db 3901 TGTGTACTGTATCATCTGAGTTGTTGTGAAAA 3936

RESULT 4  
US-10-160-719-37  
; Sequence 37, Application US/10160719  
; Publication No. US20030167528A1  
; GENERAL INFORMATION:  
; APPLICANT: Kanwarpal S. Dhugga  
; APPLICANT: Timothy G. Helentjaris  
; APPLICANT: Benjamin A. Bowen  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: thereof  
; FILE REFERENCE: 0864  
; CURRENT APPLICATION NUMBER: US/10/160,719  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/371,383A  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR APPLICATION NUMBER: US 60/096,822  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 3969  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (144)...(3401)  
US-10-160-719-37

Query Match 100.0%; Score 3936; DB 13; Length 3969;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTCCCTCGTCGGTGCAGGCGGTGGCGCGGCTCGGCGTTCCGTGAGAAACCACTCGGGG 60  
Db 1 CTTCTCCCTCGTCGGTGCAGGCGGTGGCGCGGCTCGGCGTTCCGTGAGAAACCACTCGGGG 60

QY 61 ATGAGATCTGCTGCTAGAGTGAAGAGAGCTACGCTCAGTATCCTTGCTTGTGCGCG 120  
Db 61 ATGAGATCTGCTGCTAGAGTGAAGAGAGCTACGCTCAGTATCCTTGCTTGTGCGCG 120

QY 121 GCGAAGTGAAGGAGGAGAAAGCATGAGAGCGAGCGCGGCTGTGGCGGCTCCACA 180  
Db 121 GCGAAGTGAAGGAGGAGAAAGCATGAGAGCGAGCGCGGCTGTGGCGGCTCCACA 180

QY 181 ACCGCAACGAGCTCGTCGTATCCGCCGCGAGCGCGATCCCGGCGGAGCGCGCGGG 240  
Db 181 ACCGCAACGAGCTCGTCGTATCCGCCGCGAGCGCGATCCCGGCGGAGCGCGCGGG 240

QY 241 AGCAGAACGGGCGAGGTGTCCAGATTGTGGCGCGACGACGTCCGCTTGCCCGCGGGG 300  
Db 241 AGCAGAACGGGCGAGGTGTCCAGATTGTGGCGCGACGACGTCCGCTTGCCCGCGGGG 300

QY 301 ACCCCTTGTGGCGGTGAACGAGTGCAGCTTCCCGCTGCGCGGAGCTGTGCAATACG 360  
Db 301 ACCCCTTGTGGCGGTGAACGAGTGCAGCTTCCCGCTGCGCGGAGCTGTGCAATACG 360

QY 361 AGCGCCGGGAGGCGACGAGAACTGCCCCAGTGAAGACTCGATACAAGCGCTCAAGG 420  
Db 361 AGCGCCGGGAGGCGACGAGAACTGCCCCAGTGAAGACTCGATACAAGCGCTCAAGG 420

QY 421 GCTGCCAAGTGTGACCGGTGACGAGAGAGAGAGCGGCTGATGACCTGGAACAAGT 480  
Db 421 GCTGCCAAGTGTGACCGGTGACGAGAGAGAGAGCGGCTGATGACCTGGAACAAGT 480

QY 481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCGGAGTCCATGCTTACGCGCCACATGA 540  
Db 481 TCAACTGGGACGGCCATGACTCGCAGTCTGTGGCGGAGTCCATGCTTACGCGCCACATGA 540

QY 541 GCTACGGCGGTGAGGTGACCTTAATGCGCGCGCAAGCTTTCAGCTCAACCCCAATG 600  
Db 541 GCTACGGCGGTGAGGTGACCTTAATGCGCGCGCAAGCTTTCAGCTCAACCCCAATG 600

QY 601 TTCCACTCCTCACCAACGGGCAAAATGTGATGACATCCACCGGAGCAGCAGCGCTGG 660  
Db 601 TTCCACTCCTCACCAACGGGCAAAATGTGATGACATCCACCGGAGCAGCAGCGCTGG 660

QY 661 TGCCTTCTTCATGGGTGTGGGGGAAAGAGATPACCCCTTCTTATGCGGATCCCA 720  
Db 661 TGCCTTCTTCATGGGTGTGGGGGAAAGAGATPACCCCTTCTTATGCGGATCCCA 720

QY 721 GCTTACCTGTGCAACCCAGTCTATGAGCCCATCCAAGATCTTGCTGCATATGGGTATG 780  
Db 721 GCTTACCTGTGCAACCCAGTCTATGAGCCCATCCAAGATCTTGCTGCATATGGGTATG 780

QY 781 GTAGTGTGCTTGAAGGAACGGATGGAATTGGAAGCAGAGACAAGAGAGATGCACC 840  
Db 781 GTAGTGTGCTTGAAGGAACGGATGGAATTGGAAGCAGAGACAAGAGAGATGCACC 840

QY 841 AGACGGGAGATGATGGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 900  
Db 841 AGACGGGAGATGATGGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 900

QY 901 AAGCAAGACAACAACCTGTCCAGAAATTCACCTTCATCAAGCCAGATTATCCATATA 960  
Db 901 AAGCAAGACAACAACCTGTCCAGAAATTCACCTTCATCAAGCCAGATTATCCATATA 960

QY 961 GGATGATTATCATTTATGCGCTTGTGTTTGGGGTCTTCTTCCACTACGAGTGATGC 1020  
Db 961 GGATGATTATCATTTATGCGCTTGTGTTTGGGGTCTTCTTCCACTACGAGTGATGC 1020

QY 1021 ATCCGGTGAATGATGATTTGCTTTGTGGCTCATATCTGTTATCTGTGAATCTGTTG 1080  
Db 1021 ATCCGGTGAATGATGATTTGCTTTGTGGCTCATATCTGTTATCTGTGAATCTGTTG 1080

QY 1081 CCATGCTTGGATTCTTGATCAATTCCAAAGTGTTCCTTAFTGAGAGAGACTTACC 1140  
Db 1081 CCATGCTTGGATTCTTGATCAATTCCAAAGTGTTCCTTAFTGAGAGAGACTTACC 1140

QY 1141 TAGACCGGCTGCTACTGAGTTGCAAGGAAGGCCAGCATCTCACTTGTCCCAATTG 1200  
Db 1141 TAGACCGGCTGCTACTGAGTTGCAAGGAAGGCCAGCATCTCACTTGTCCCAATTG 1200

QY 1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGAACCTCCTTTGGTCAACAACAATACTG 1260  
Db 1201 ATTTCTTTGTCAGTACGGTTGATCCCTTAAAGAACCTCCTTTGGTCAACAACAATACTG 1260

QY 1261 TTCTATCTATCCTTTGCGTGATTAATCCTGTGATTAAGTTTCTGTATGTTCTGATG 1320  
Db 1261 TTCTATCTATCCTTTGCGTGATTAATCCTGTGATTAAGTTTCTGTATGTTCTGATG 1320

QY 1321 ATGTGCTCAATGCTTAACGTTTGAACATTATCTGAACACTGTAATTTGCAAGAAT 1380  
Db 1321 ATGTGCTCAATGCTTAACGTTTGAACATTATCTGAACACTGTAATTTGCAAGAAT 1380

QY 1381 GGGTCTCTTCTGCAACCGGTACAAATTTGAACCTCGCGCTCAGAGTGTACTTCCAA 1440  
Db 1381 GGGTCTCTTCTGCAACCGGTACAAATTTGAACCTCGCGCTCAGAGTGTACTTCCAA 1440

QY 1441 AGAAGTAACTACTTGAAGAACAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAA 1500  
Db 1441 AGAAGTAACTACTTGAAGAACAAGTGGCAGCAAACTTTGTTAGGAGAGAGAGCAA 1500

QY 1501 TGAAGAGAGATATGAGGAATTCAGGTGAGAAATCAATGCTTGAATTTGCCAAAGCCAGA 1560  
Db 1501 TGAAGAGAGATATGAGGAATTCAGGTGAGAAATCAATGCTTGAATTTGCCAAAGCCAGA 1560

QY 1561 AAGTTCCTGAAGAAAGATGGAACAATGCAAGATGAAACCCCTGGCTGGAACAATGTT 1620  
Db 1561 AAGTTCCTGAAGAAAGATGGAACAATGCAAGATGAAACCCCTGGCTGGAACAATGTT 1620

QY 1621 GTGATCATCTTGAATGATTCAGTCTTCTTGCCAAAGCGAGGCTTGACTGTGAGG 1680  
Db 1621 GTGATCATCTTGAATGATTCAGTCTTCTTGCCAAAGCGAGGCTTGACTGTGAGG 1680

QY 1681 GAAATGAACCTGCCACGATTGTTATGTTTCTAGAGAGAAACGACGCTATTAACCATC 1740  
Db 1681 GAAATGAACCTGCCACGATTGTTATGTTTCTAGAGAGAAACGACGCTATTAACCATC 1740

QY 1741 ATAAGAAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
Db 1741 ATAAGAAAGCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1800

QY 1801 CATATTTGTTAAACCTTGATTTGATCACTACATCAACAACGAGCTATTAAGGAAG 1860  
Db 1801 CATATTTGTTAAACCTTGATTTGATCACTACATCAACAACGAGCTATTAAGGAAG 1860

QY 1861 CAATGTTTTATGATGAGCCCTTAACTAGAAAGAGGTTGCTATGTAAGTTCCTC 1920  
Db 1861 CAATGTTTTATGATGAGCCCTTAACTAGAAAGAGGTTGCTATGTAAGTTCCTC 1920

QY 1921 AAAGATTGATGGATGATGCGCATGACCGATATGCTAACCGGAATGTTCTTTT 1980  
Db 1921 AAAGATTGATGGATGATGCGCATGACCGATATGCTAACCGGAATGTTCTTTT 1980

QY 1981 ATATCAACATGAAGGTTTGATGTTATTCAGGGTCCAATTTATGTTGTTACTGATG 2040  
Db 1981 ATATCAACATGAAGGTTTGATGTTATTCAGGGTCCAATTTATGTTGTTACTGATG 2040

QY 2041 TATTTAGAGGCGAGGCTTATATGTTATGATGCCCCCAAAACAAGAGCCACCATCA 2100  
Db 2041 TATTTAGAGGCGAGGCTTATATGTTATGATGCCCCCAAAACAAGAGCCACCATCA 2100

QY 2101 GGACTTGCAACTGCTGCCCCAAGTGTGCTTTTGTGCTGTGCTTGGCAATAGGAAGC 2160



Db 2101 GGACTTGCAACTGCTGGCCCAAGTGGCTTTTGCTGTGCTGCTTGCGCAATAGAGAGC 2160  
Qy 2161 AAAAGAGACTACCAAAACCCAAAACAGAGAAAAAGTTATTTTTCAGAAAGAG 2220  
Db 2161 AAAAGAGACTACCAAAACCCAAAACAGAGAAAAAGTTATTTTTCAGAAAGAG 2220  
Qy 2221 AGAACCAATCCCTGTCATATGCTCTTGGTGAATTTGACGAAGCTCTCCAGAGCTGAGA 2280  
Db 2221 AGAACCAATCCCTGTCATATGCTCTTGGTGAATTTGACGAAGCTCTCCAGAGCTGAGA 2280  
Qy 2281 ATGAAAAAGCCGGTATTGTAATCAAAAAATTAGAAAAAGAAATTGGCCAATCTTCTG 2340  
Db 2281 ATGAAAAAGCCGGTATTGTAATCAAAAAATTAGAAAAAGAAATTGGCCAATCTTCTG 2340  
Qy 2341 TTTTGTACATCCACACTTCTCGAATGGTGGAACTTGAAGAGTGCAGTCCCTGCTT 2400  
Db 2341 TTTTGTACATCCACACTTCTCGAATGGTGGAACTTGAAGAGTGCAGTCCCTGCTT 2400  
Qy 2401 CTCTTTGAAGAAGCTATACATGTCATTAAGTTGTTGTTATGAAGACAAGACAGACTGGG 2460  
Db 2401 CTCTTTGAAGAAGCTATACATGTCATTAAGTTGTTGTTATGAAGACAAGACAGACTGGG 2460  
Qy 2461 GAAAAGAGATTGGCTGGATCTATGATCAGTTACAGAAAGATATTCTAACTGTTCAAGA 2520  
Db 2461 GAAAAGAGATTGGCTGGATCTATGATCAGTTACAGAAAGATATTCTAACTGTTCAAGA 2520  
Qy 2521 TGCAATGTCATGGTGGCGGTCAATTTACTGCATACCTAAACGGGTGCATTCAAAAGTT 2580  
Db 2521 TGCAATGTCATGGTGGCGGTCAATTTACTGCATACCTAAACGGGTGCATTCAAAAGTT 2580  
Qy 2581 CTGCACTCTGAATCTTTCAGATCGTCTTCAACCAAGTCTCGGTGGGCTCTGGGTCTA 2640  
Db 2581 CTGCACTCTGAATCTTTCAGATCGTCTTCAACCAAGTCTCGGTGGGCTCTGGGTCTA 2640  
Qy 2641 TTGAGATCTTCTTCAGCAATCATGCCCCCTTTGGTATGGGTATGGTGGCGGTCTGAAT 2700  
Db 2641 TTGAGATCTTCTTCAGCAATCATGCCCCCTTTGGTATGGGTATGGTGGCGGTCTGAAT 2700  
Qy 2701 TTTTGGAAAAAGATTTTCTCTACATCAACTCCATCGTGTATCCTTGGACATCTATCCCTCT 2760  
Db 2701 TTTTGGAAAAAGATTTTCTCTACATCAACTCCATCGTGTATCCTTGGACATCTATCCCTCT 2760  
Qy 2761 TGGCTTACTGTACATGCGCTGCGCATCTGTTTATTGACAGGAAATTATCACTCCAGAGC 2820  
Db 2761 TGGCTTACTGTACATGCGCTGCGCATCTGTTTATTGACAGGAAATTATCACTCCAGAGC 2820  
Qy 2821 TGAATATATGTGCGCAGCTGTGGTCAATGTCATTTTATCTGCATTTTGTCTACGAGCA 2880  
Db 2821 TGAATATATGTGCGCAGCTGTGGTCAATGTCATTTTATCTGCATTTTGTCTACGAGCA 2880  
Qy 2881 TCCTAGAAAAATGAGATGAGTGGTGTGGAATTGATGACTGGTGAAGAAATGAGCAGTTCT 2940  
Db 2881 TCCTAGAAAAATGAGATGAGTGGTGTGGAATTGATGACTGGTGAAGAAATGAGCAGTTCT 2940  
Qy 2941 GGGTCATTGGAGGTGTGTCTTCAACCTCTTGTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000  
Db 2941 GGGTCATTGGAGGTGTGTCTTCAACCTCTTGTGCTGTGTTCCAGGGACTTCTCAAGGTCA 3000  
Qy 3001 TAGCTGCTGTGATACAAAGCTTCAACCGTGACATCAAAAGGGTGAGATGAGAGATTCT 3060  
Db 3001 TAGCTGCTGTGATACAAAGCTTCAACCGTGACATCAAAAGGGTGAGATGAGAGATTCT 3060  
Qy 3061 CAGAGCTATATACATTCAAAATGAGTACCTTATTGATACCTCTACCACTTGTCTTCTAT 3120  
Db 3061 CAGAGCTATATACATTCAAAATGAGTACCTTATTGATACCTCTACCACTTGTCTTCTAT 3120  
Qy 3121 TGAACCTCATTTGGTGTGCTGCGCTTTCAAATGCGATCAATAACGATATGAGTCAAT 3180  
Db 3121 TGAACCTCATTTGGTGTGCTGCGCTTTCAAATGCGATCAATAACGATATGAGTCAAT 3180  
Qy 3181 GGGGCCCCCTCTTTGGAAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTATCCCT 3240  
Db 3181 GGGGCCCCCTCTTTGGAAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTATCCCT 3240

Db 3181 GGGGCCCCCTCTTTGGAAAGCTATTCTTTGCATTTTGGGTGATTGTCCATCTTATCCCT 3240  
Qy 3241 TTCTCAAAAGTTTGGTGAAGGCCAAAACAGACACCAAGATTGTCACTGTGTGCCA 3300  
Db 3241 TTCTCAAAAGTTTGGTGAAGGCCAAAACAGACACCAAGATTGTCACTGTGTGCCA 3300  
Qy 3301 TTCTGCTGCTTCAATCTTCTCGCTCTTGGTGGTTCGATGATCCTTCTTGGCAAGG 3360  
Db 3301 TTCTGCTGCTTCAATCTTCTCGCTCTTGGTGGTTCGATGATCCTTCTTGGCAAGG 3360  
Qy 3361 ATGATGCTCCGCTTCTTGAAGAGTGTGTTGGATTGCACTAGGATGTCAATGATCAG 3420  
Db 3361 ATGATGCTCCGCTTCTTGAAGAGTGTGTTGGATTGCACTAGGATGTCAATGATCAG 3420  
Qy 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGTGTGTTGCCCATATTCAGTGT 3480  
Db 3421 CTCCCCCAATCTGCATATGCTTGAAGTATATTTCTGTGTGTTGCCCATATTCAGTGT 3480  
Qy 3481 CTGTAGATAAGACATGAATAATGCCAAGTTCTTTGATCCATGRTGAACCTTAA 3540  
Db 3481 CTGTAGATAAGACATGAATAATGCCAAGTTCTTTGATCCATGRTGAACCTTAA 3540  
Qy 3541 TATCTGAGAGATATACCTGGGGGAAAAATGAGGCTGGCGCAATCCTGTGAGTTGGCGG 3600  
Db 3541 TATCTGAGAGATATACCTGGGGGAAAAATGAGGCTGGCGCAATCCTGTGAGTTGGCGG 3600  
Qy 3601 TGAATATACAGCATATGCAAGTGTGATTGTGTCAGCAATCTTATTAATTGTTGCGCAATA 3660  
Db 3601 TGAATATACAGCATATGCAAGTGTGATTGTGTCAGCAATCTTATTAATTGTTGCGCAATA 3660  
Qy 3661 TAGATGGGCTGAGCCGAACAGCAAGGATTTTGAATCTGCACCTGCTCCCGTGAACAACT 3720  
Db 3661 TAGATGGGCTGAGCCGAACAGCAAGGATTTTGAATCTGCACCTGCTCCCGTGAACAACT 3720  
Qy 3721 TGGTCTCATTAAGGACGAGGAATGCATCTGCCAGTGAACAGAGCAACCTGCACATT 3780  
Db 3721 TGGTCTCATTAAGGACGAGGAATGCATCTGCCAGTGAACAGAGCAACCTGCACATT 3780  
Qy 3781 ATTTATGTATGCTGTTCAATTGAGGGCTTGTTCATTAATGTTCTGTCTATAGAAAA 3840  
Db 3781 ATTTATGTATGCTGTTCAATTGAGGGCTTGTTCATTAATGTTCTGTCTATAGAAAA 3840  
Qy 3841 AACAGATATATGACATTAATCTATAGTTAATTAAGTATGTAATGCGCTGTTTGTGT 3900  
Db 3841 AACAGATATATGACATTAATCTATAGTTAATTAAGTATGTAATGCGCTGTTTGTGT 3900  
Qy 3901 TGTGTACTGTATCATCTGAGTTGCTTTGTGAAAA 3936  
Db 3901 TGTGTACTGTATCATCTGAGTTGCTTTGTGAAAA 3936

RESULT 5  
US-09-900-237-3  
: Sequence 3, Application US/09900237  
: Patent No. US20020120124A1  
: GENERAL INFORMATION:  
: APPLICANT: Allen, Stephen  
: TITLE OF INVENTION: Plant Cellulose Synthases  
: FILE REFERENCE: BB1170 US CIP  
: CURRENT APPLICATION NUMBER: US/09/900,237  
: PRIORITY FILING DATE: 2001-07-06  
: PRIOR APPLICATION NUMBER: 60/092,844  
: PRIOR FILING DATE: 1998-07-14  
: PRIOR APPLICATION NUMBER: PCT/US99/15871  
: PRIOR FILING DATE: 1999-07-13  
: PRIOR APPLICATION NUMBER: 09/720383  
: NUMBER OF SEQ ID NOS: 33  
: SOFTWARE: Microsoft Office 97  
: SEQ ID NO 3  
: LENGTH: 3776  
: TYPE: DNA  
: ORGANISM: Zea mays

US-09-900-237-3

```
Query Match      65.9%; Score 2592.6; DB 10; Length 3776;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 2885; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

QY 104 CTCTGCCTTCGTCCGCGCGGAGTGAAGGGAGGAAGCGATGAGCGGAGCGCGGCT 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 CGCTTGCCCCGGGATCTGGAGATCTGGTAGCGCCAGGGGATGAGGCCAGCGCGGCT 198

QY 164 GGTGGCCGGCTCCCAACAACCGCAACGAGCTGCTCATCCGCCCGCAACGGGATCCCG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 GGTCCCGGCTCGCAACAACCGAACGAGCTGCTCATCCGCCCGCATGGCGAGCCAGG 258

QY 224 GCCGAAGCCCGCGCGGAGCAGAACCGGAGGTGCGCAGATTGCGCGCAGACGTCGG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 GCCGAAGCCCATGGAACGAGCGGAACGGCCAGGTGCGCAGATTGCGCGCAGACGTCGG 318

QY 284 CCTTGCCCCGCGGAGACCCCTTCGTGCGGTGCAACGAGTGCCTTCCCGTGCCTG 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 GCGCAACCCCGACGGGAGCGCTTCGTGCGCTGCAACGAGTGCCTTCCCATCTGCCG 378

QY 344 GGACTGCTACGATAACGAGCGCCGAGGAGCAGCAACTGCCCCAGTGCAGAAGACTCG 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 GGACTGCTACGATACGAGCGCGCGGAGGCAAGCAACTGCCCCAGTGCAGAAGACCCG 438

QY 404 ATACAAGCGCTCTCAAGGGGCTGCCAACGTGTGACCGGTGACGAGGAGAGGACGGCGTGA 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 CTTCAGCGCTCAAGGGGCTCGCGCGCGCTGCGCGGAGCAGAGGAGAGGACGGCGTGA 498

QY 464 TGACCTGACACAAGAGTTCAACTGG--GACGGCCATGACTCGCAGTGTGTGCCGAGTC 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 CGACCTGGAGAACGAGTTCAACTGGAGCGACAGCAAGACTCCAGTACTCGCGGAGTC 558

QY 521 CATGCTCTACGGCCACATGAGCTACGGCGGTGAGGTGAGCCCTAATGCGCGCCACAAGC 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 CATGCTCCACGGCCACATGAGCTACGGCGCGCGCGACCTCGACGGCGTCCCGCAGCC 618

QY 581 TTTCCAGCTCAACCCCAATGTTCCACTCTCTCAACCAAGGCAAAATGTTGATGACATCCC 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 ATTCACACCCCATCCCAATGTTCCCTCCCTCAACCAAGGACAGATGTTGATGACATCCC 678

QY 641 ACCGAGCAGCAGCGCGCTGTGCTTCTTCAATGGGTGTGGGGGAAGAGATACATCC 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 GCCGGACAGCAGCGCTTGTGCTCTCTGTTGGGTGGCGGGGGAAGAGATTCAACC 738

QY 701 CTTTCTTATGGGATCCAGCTTACCTGTGCAACCCAGGTCTATGACCATCCAAGGA 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 TCTCCCGTACCGGATCCCAACCTTCTGTGCAACGAGGTCTATGACCCTTCCAAGGA 798

QY 761 TCTTGCTGATATGGGTATGTTAGTGTGCTTGAAGAAAGGATGAGATGGAAGCA 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 TCTGCGCGCATATGGCTACGGGAGCTAGCATGGAAGAGAGATGAGAGCTGGAAGCA 858

QY 821 GAGACAAGAGAGATGACACGAGCGGGAATGATGGTGTGATGATGATGATGATGC 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 GAAGCAGAGAGAGATGACACGAGAGAGAGATGGCGGCGGATGATGATGATGATGC 918

QY 881 TGATCTACCACTAATGATGAAGCAAGACAACACTGTCCAGGAAATTCACCTTCATC 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 919 AGATCTACCACTAATGATGAAGCTAGACAGCCAATTGTCCAGAAAGATCCCGCTTC 978

QY 941 AAGCCAGATTATCCATATAGGATGATTATATTCGGCTGTGTTTGGGGTCTT 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 AAGCCAAATCAACCCCTATAGGATGATTATTAATTCGGCTAGTGTGTTTGTCTT 1038

QY 1001 CTTCACCTACCGAGTATGATCGGCTGAATGATGATTGCTTGTGCTCATATCTGT 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1039 CTTCACCTACCGAGTATGATCGGCTGTGATGATGATTGCTTATGCTCATATCTGT 1098

QY 1061 TATCTGGAATCTGTTGTCATGTTGATCTTGATCTTGATCAATTCCAAAGTGTGCC 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1099 GATCTGTGAATTTGTTGTCATGTTGATCTTGATCTTGATCAAGTTTCCAAGTGTGCC 1158
```

```
QY 1121 TATTGAGAGAGACTTACTAGACCGGCTGTCACTGAGGTTGACAAGAGGCCAGCC 1180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1159 TATCGAGAGGGAACCTATCTTGACCGGCTGAGTTTAAGTTTGACAAGAGAGGCACTC 1218

QY 1181 ATCTCAACTTGCTCCAATTGATTTCTTTGTGAGTACGGTTGATCCCTTAAGGAACCTCC 1240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1219 TTCTCAACTCGCCCTGTTGATTTCTTTGTGAGTACGGTTGATCCCTTAAGGAACCTCC 1278

QY 1241 TTTGCTACAACAATACTGTTCTATCTATCTTTCGTTGGAATTATCTGTTGATAAGT 1300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1279 ATTGCTACTGCTAATACTGTTCTATCTATCTTTCGTTGGAATTATCCAGTTGATAAGT 1338

QY 1301 TTCTTGCTATGTTTCTGATGATGCTGCAATGCTAAAGTTGAAGCATTATGAAAC 1360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1339 TTCAATGCTACGTTTCTGATGATGCTGCTGCATGCTGACATTTGAAGCATTTGTAAC 1398

QY 1361 ATCTGAATTTGCAAGAATGGCTTCTTCTGCAACGGTACAATATTTGAACCTCGCGC 1420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1399 ATCTGAATTTGCAAGAATGGCTTCTTCTGCAAAAGATATAAGCTTGAGCCTCGTGC 1458

QY 1421 TCCAGAGTGTACTTCCAAACAGAAATAGACTACTTGAAGAACAAGGTGGCAGCAACTT 1480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1459 TCCAGAGTGTACTTCCAAACAGAAATAGACTACTTGAAGAACAAGGTGGCAGCAACTT 1518

QY 1481 TGTTAAGGAGAGAGAGCAATGAAGAGAGATGAGAAATCAAGGTGAGATCAATGC 1540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1519 TGTTAAGAAACGAGAGCAATGAAGAGAGATGAGAAATCAAGGTGAGATCAATGC 1578

QY 1541 CTTAGTTCCAAAGCCCAAGAGTCTCTGAAGAAGGATGCAATGCAAGATGAACCC 1600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1579 CTTGTTGCTAAAGCCCAAGAGTCTCTGAAGAAGGATGCAATGCAAGATGAACCTCC 1638

QY 1601 CTGCTTGAAACAATGTTCTGATTCATCTCGAATGATTCAGTCTTCTTGCCAAAG 1660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1639 ATGCGCCGGAATAATGTCCGATCATCTCGAATGATTCAGTCTTCTTGCTCAAG 1698

QY 1661 CGGAGCCTTGACTGTGAGGGAATGAACCTGCCACGATTTGTTATGTTCTAGAGAA 1720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1699 TGGTGGCATGATGTGGAAGGAATGAAGCTGCCCTCGATTGTTATGTTCAAGAGAA 1758

QY 1721 ACGACAGGCTATAACCATATTAAGAAAGCTGTGCTATGATGATTTGTCGAGTCTC 1780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1759 ACGGCAAGCTACAACCATACAGAAGGCTGTGCTATGATGATGATGATCCGAGTCTC 1818

QY 1781 TGCTGTACTAACAAATGCTCCATTTGTTAACTTGATTTGATCACTACATCAACAA 1840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1819 TGCTGTACTAACAAATGCTCCATTTGTTGATGATGATGATGATGATGATCAATAA 1878

QY 1841 CAGCAGGCTATAAAGGAAGCAATGTTTATGATGATGATGATGATGATGATGAT 1900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1879 TAGTAAGGCTATAAAGGAAGCAATGTTTATGATGATGATGATGATGATGATGAT 1938

QY 1901 TTGCTATGTACAGTTCCTCTCAAGATTTGATGGATTGATGCCATGACCGATATGCTAA 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1939 TTGCTATGTACAGTTCCTCTCAAGATTTGATGGATTGATGCCATGATGATGCTAA 1998

QY 1961 CCGGATGTTGCTTTTGTGATCAACATGAAAGTTTGAATGATTCAGGGTCCAAT 2020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1999 CAGAAATGTTGCTTTTGTGATCAACATGAAAGTTTGAATGATTCAGGGGCCAAT 2058

QY 2021 TTATGTTGTACTGAGTGTATTTAGAAAGGAGGATTAATGTTATGATGCCCCCAA 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2059 TTATGTTGTACTGAGTGTATTTAGAAAGGAGGATTAATGTTATGATGCCCCCAA 2118

QY 2081 AACAAAGAACCCATCAAGACTTGCAACTGCTGCGCAAGTGTGCTTTGTGCTGTTG 2140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2119 AACAAAGAACCCATCAAGACTTGCAACTGCTGCGCAAGTGTGCTTTGTGCTGTTG 2178

QY 2141 CTGCTTGCAATAGGAAGCAAAAGAACTACCAAAACCAAAACAGAGA-----AGAA 2194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2179 CTGCTTGCAATAGGAAGCAAAAGAACTACCAAAACCAAAACCAAACTTAATTTGAGAA 2238
```

QY	2195	AAAGTATTATTTTTCAGAAAGAGAACCAATCCCCTGCATATGCTCTTGGTGAAT	2254
Db	2239	GATAAAGAACTTTTAAAGAAAAAGAAATCAAGCCCTGCATATGCTCTTGGAAT	2298
QY	2255	TGACGAAGCTGCTCCAGAGCTGAGATGMAAAGCCGGTATTGTAAATCAACAAAAATT	2314
Db	2299	TGATGAAGCCGCTCCAGAGCTGAAATGMAAAGGCTAGTATTGTAATCAACAGAGTT	2358
QY	2315	AGAAAAGAATTTGGCCAATCTTCTGTTTTGTACATCCACACTTCTCGAATGGTGG	2374
Db	2359	GGAAAAGAATTTGGCCAGTCTTCAGTTTTTGTTCATCCACACTTCTTGAGATGCTGG	2418
QY	2375	AACCTTGAAGAGTGCAGAGTCTGCTTCTCTTTGMAAGAACTATACATGTCATTAGTTG	2434
Db	2419	AACCTGAAGAGTGCAGTCCAGCTTCTCTTGAAGAACTATACATGTCATCAGTTG	2478
QY	2435	TGGTTATGAAGACAAGACAGACTGGGGAAAAAGAGATTGGCTGGATCTATGATCAGTTAC	2494
Db	2479	TGATATGAAGACAAGACAGGCTGGGGAAAAAGATATTGGTTGATTATGATCAGTCAC	2538
QY	2495	AGAAGATATTCTAAGTGTTCAGATGCATTGTCAATGTTGGCGGTCAATTTACTGCAT	2554
Db	2539	AGAAGATATTCTTACTGGGTTTAAGATGCAGTCCATGGTTGGCGGTCAATTTACTGCAT	2598
QY	2555	ACCTAAACGGGTGTCATTCAAAGGTTCTGCACCTCTGAAATCTTTCAGATCGTCTTCA	2614
Db	2599	ACCTAAACGGGCGCCCTTCAAAGGTTCCGACCTCTCAATCTTCCGATCGTCTTCA	2658
QY	2615	GGTGCTTCGGTGGGCTCTTGGGTTCTATTGAGATCTTCTTCAGCAATCATTTGCCCTTTG	2674
Db	2659	GGTCTTCGGTGGGCTCTTGGTTCAATTGAAATTTCTTCAGCAACCACTGCCCTCTCTG	2718
QY	2675	GATGGGTATGGTGGCGGTCTGAATTTTGGAAAGATTTTCTACATCACTCCATCGT	2734
Db	2719	GATGGGTATGGTGGTGAATAAGTTCCTGMAAGGTTTTCGTACATTAATCTCCATCGT	2778
QY	2735	GTATCCTTGGACATCTATTCCTCTTGGCTTACTGTACATTTGCCATCTGTTTATT	2794
Db	2779	ATACCTTGACATCTATCCGCTCTTGCCCTATTGACATTTGCCATCTGCTTGTCT	2838
QY	2795	GACAGGAAATTTATCACTCCAGAGCTGAAATATGTTGCCAGCTGTGTTCAATGCACT	2854
Db	2839	GACAGGAAATTTATCAAGCAGAGCTTAACAATGTTGCCAGCTGTGTTCAATGCACT	2898
QY	2855	TTTTATCTGCATTTTGTCTACGAGCATCCTAGAAATGAGATGAGTGTGTAATTGA	2914
Db	2899	TTTCATCTGCATTTTGTCTACGAGCATCCTGAAATGAGATGAGTGTGTAAGCATGA	2958
QY	2915	TGACTGTTGAGGAATGAGAGTCTTGCGTCATTGGAGGTGTCTCACACCTCTTTGC	2974
Db	2959	TGACTGTTGAGAAACGAGCAGTTTGGGTCAATTGAGCGGTGCTTCACATCTCTTGC	3018
QY	2975	TGTGTTCCAGGACTTCTCAAGTTCATAGCTGATGTTGATACAAGCTTCAACGTCATC	3034
Db	3019	TGTGTTCCAGGACTCCTCAAGTTCATAGCTGATGATACAGAGCTTCACTGTGATC	3078
QY	3035	AAAGGTGAGATGATGAGAGTCTCAGAGCTATATACATTCAATGGACTACTTAATT	3094
Db	3079	CAAGGCGGAGACGACGAGAGTCTCAGAGCTGTACATTCAATGGAAGCCTTCT	3138
QY	3095	GATACCTCTTACCACTTGTCTTATTTGAACCTTCAATGGTGTGCTGCGCTTCAAA	3154
Db	3139	GATACCTCCGACCACTGTCTCTTCTGAACCTTCAATGAGTGTAGCTGTCATCTCAA	3198
QY	3155	TGCGATCAATTAACGATATGAGTCAATGGGCCCCCTTTGGGAAGCTATTCTTGCAAT	3214
Db	3199	TGCGATCAACACGATATGATCATGGGGCCCCCTGTTGGGAAGCTTTCTTGCAAT	3258
QY	3215	TTGGGTGATGTCCATCTTTATCCCTTTCTCAAAAGTTGGTTGGAAGGCAAAACAGAC	3274
Db	3259	TTGGGTGATGTCCATCTTTATCCCGTTCTCAAGGGTCTGGTTGGAGGCAAGACAGAC	3318
QY	3275	ACCAACGATGTGATCGTCTGCTCATTCTGCTGGCTTCAATCTTCTGCTCTTTGGGT	3334

```

Db      3319  GCCAACGATGTGTCATGTGTGCTGCATCCCTCGGCTTCGATCTTCTCGCTGCTTGGGT 3378
QY      3335  TCGGATTGATCCTTTCCTTGGCAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTGGA 3394
Db      3379  CCGGATCGACCCCTTCTTGGCGAAGGATGATGTGTCCCTGTTGGAGGAGTGTGTTGGA 3438
QY      3395  TTGCACTAGAGTGTGATGCATCAGCTCCCCC 3427
Db      3439  TTGCACTAGAGGTGACGACGTGACTTCCCC 3471

RESULT 6
US-10-209-059-41
; Sequence 41, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3487
; OTHER INFORMATION: n = A,T,C or G
US-10-209-059-41

Query Match      64.7%; Score 2546.8; DB 13; Length 3568;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 2829; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

QY      177  CACAACCGCAACGAGCTGTGTCATCCGCGCGAAGCGGATCCCGGGCCGAAGCCGCG 236
Db      6    CCCACGCGTCCGAGCTGTGTCATCCGCGCGAAGCGGATCCCGGGCCGAAGCCATG 65
QY      237  CGGAGCAGAACGGCGCAGGTGTGCCAGATTTGCGCGCAGCAGCTGCGCCTTGCCCCGCG 296
Db      66  GACCAGCGGAACGGCGCAGGTGTGCCAGATTTGCGCGCAGCAGCTGCGCGCAACCCGAC 125
QY      297  GGGGACCCCTTCTGTGCGCTGCAACGAGTGCCTTCCCCGCTCTGCCGGGACTGCTACGAA 356
Db      126  GGGGAGCCTTTCGTGCGCTGCAACGAGTGCCTTCCCCATCTGCCGGGACTGCTACGAG 185
QY      357  TACGAGCGCGGGAGGGGACGCAAGACTGCCCCCAGTGCAAGACTCGATACAAGCGCTC 416
Db      186  TACGAGCGCGGGAGGGGACGCAAGACTGCCCCCAGTGCAAGACCCGCTTCAAGCGCTTC 245
QY      417  AAGGCGTGCACAAGTGTGACCGGTGACGAGAGGAGGACGGCGTGCATGACCTGAGAAC 476
Db      246  AAGGCGTGCAGCGCGCGTGCAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 305
QY      477  GAGTTCAACTGG--GACGGCCATGACTCGCAGTGTGTGGCCGAGTCCATGCTTACGCG 533
Db      306  GAGTTCAACTGGAGCGACACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 365
QY      534  CACATGAGCTACGGCGCGTGAAGGTGACCTTAATGCGCGGCCACAAGCTTTCAAGCTCAAC 593
Db      366  CACATGAGCTACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 625

```



[illegible]

QY	1674	TGTGAGGGAATGAACTGCCACGATTGGTTATGTCTTAGAGAGAAACGACGAGCTAT	1733
Db	1506	GTGGAAGGAATGAGCTGCTCGATTGGTTATGTTTCAAGAGAAAAACGCCAGGCTAC	1565
QY	1734	AACCATCAATAAGAAAGCTGFGCTATGAATGCAATGGTCCGAGTCTCTGCTGTACTAACA	1793
Db	1566	AACCATCAACAAGAGGCTGTGCTATGAATGCAATGGTCCGAGTCTCTGCTGTACTAACT	1625
QY	1794	AATGCTCCATATTTGTTAACTTGATTTGTGATCACTACATCAACAACAGCAAGGCTATA	1853
Db	1626	AATGCTCCTTATTTGCTGAACCTTGATTTGTGATCACTATATCAATATAGTAAGCTATA	1685
QY	1854	AAGGAAGCAATGTGTTTTATGATGAGCCCTTTACTAGGAAGAAGGTTTGTATGTACAG	1913
Db	1686	AAGGAAGCAATGTGTTTTATGATGAGTCCCTTTGCTTTGGAAAGAAGTTTGTATGTACAG	1745
QY	1914	TTCCCTCAAAAGATTGATGGGATTGATCGCCATGACCGATATGTCTAACCGGAATGTTGC	1973
Db	1746	TTTCCTCAAAAGATTGATGGGATTGATCGCCATGATGATATGCTAACAGAAATGTTGTC	1805
QY	1974	TTTTTTGATATCAACATGAAGAGTTTGATGTGTAATTCAGGGTCCAAATTTATGTGTAAT	2033
Db	1806	TTTTTTCGATATCAACATGAAGAGTTTGATGTGTAATTCAGGGGCCAATTTATGTGGGTAAT	1865
QY	2034	GGATGTGATTTAGAAGGACGAGCATTTATATGTTATGATGCCCCCAAAACAAGAACCA	2093
Db	1866	GGATGTGCTTCAGAAAGCAGGCAATTTATGTGCTACGATGCTCCCAAAACAAGAACCA	1925
QY	2094	CCATCAAGGACTTGCACTGCTGGCCCAAGTGGTCTTTGCTGTTGCTGCTTTGGCAAT	2153
Db	1926	CCATCAAGAACTTGCACTGCTGGCCAAAGTGGTCAATTTGCTGTTGCTGTTTGGTAAC	1985
QY	2154	AGGAAGCAAAAGAGACTACCAAAACCCAAACACAGAGA-----AGAAAAAGTTATTAATT	2207
Db	1986	AGGAAGACCAAGAGAGAGACCAAGACCTTAACCTTAATTGAGAGATTAAGAAACTT	2045
QY	2208	TTCAAGAAAGAGAGAACCAATCCCCCTGCATATGCTTTGTTGAATTGACGAAGCTGCT	2267
Db	2046	TTTAAGAAAAAGAAATCAAGCCCCCTGCATATGCTTTGTTGAATTGATGAAGCCGCT	2105
QY	2268	CCAGAGCTGAGAATGAAAAAGCCGCTATTGTAAATCAACAAAAATTAGAAAAAGAAATTT	2327
Db	2106	CCAGAGCTGAAAAATGAAAAAGGCTAGTATTGTAAATCAACAGAAGTTGAAAAAGAAATTT	2165
QY	2328	GGCCAACTCTCTGTTTTTTGTACATCCACACTTCGAGAAATGTTGGAACCTTGAAGAT	2387
Db	2166	GGCCAGCTTTCAGTTTTTTGTTCATCCACACTTCGAGAAATGTTGGAACCTTGAAGAT	2225
QY	2388	GCAAGTCCCTGCTTCTTTTGAAGAAAGCTATATCATGTCTTAAGTTGTGTGTTATGAAGAC	2447
Db	2226	GCCAGTCCAGCTTCTCTCTGAAAGAAAGCTATATCATGTCTTAAGTTGTGTGTTATGAAGAC	2285
QY	2448	AAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCAAGTTACAGAAAGATATTCTA	2507
Db	2286	AAAAACAGGCTGGGAAAAAGATATTGGTTGATTTATGATCACTACAGAAAGATATTCTT	2345
QY	2508	ACTGGTTTCAAGATGCAATGTCTATGTTGCGGTCAATTTACTGCAATACCCTAAACGGGTT	2567
Db	2346	ACTGGTTTAAAGATGCACTGCCATGTTGGCGGTCAATTTACTGCATACCTAAACGGGCC	2405
QY	2568	GCATTCAAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAACCAAGGTGCTTCGGTGG	2627
Db	2406	GCCTTCAAAAGGTTCCGCACCTCTCAATCTTTCGATCGTTTTTCAACCAAGGTCTTCGGTGG	2465
QY	2628	GCTCTTGGGTCTATTGAGATCTTCTTCAGCAATCATTGGCCCTCTTTGGTATGGGTATGGT	2687
Db	2466	GCTCTTGGTTCATTTGAATTTTGTTCAGCAACCACTGCCCCCTCTGTGATATGGGTATGGT	2525
QY	2688	GGCGGCTCTGAAATTTTGGAAAGATTTTCTCTACATCAACTCCATCGTGTATCTTTGGACA	2747
Db	2526	GGTGGACTAAAGTTCCTGGAAAGGTTTTTCGTACATTAACCTCCATCGTATACCTTTGGACA	2585
QY	2748	TCTATTTCCCTCTTGGCTTACTGTACATTTGCCCTGCCACTCTGTTTATTTAGACAGGAAATTT	2807

```
Db      2586 TCTATCCCGCTCTTGGCCTATTGACATTGCCCTGCCATCTGCTTGTCAGAGGAATTT
      Oy      2808 ATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGCTCATGTCACTTTTATCTGCATT
      Db      2646 ATCAGCCAGAGCTTAACAATGTTGCCAGCCTCTGTTCACTGTCATCTTTTCATCTGCATT
      Oy      2868 TTGTCTAGAGCATCTTAGAATGAGTGAAGTGTGTTGGAATGATGACTGTGAGG
      Db      2706 TTGTCTAGAGCATCTTAGAATGAGTGAAGTGTGTTGGAATGATGACTGTGAGG
      Oy      2928 AATGAGAGTTCTTGGGTCAATTGAGGTGTCTCACAACCTCTTGTCTGTCCAGGA
      Db      2766 AACGAGAGTTTGGGTCAATTGAGGCGGTCTTCAACATCTTGTCTGTCCAGGA
      Oy      2988 CTCTCAAGTTCATAGCTGTGTGATACAAGCTTCAACCGTGACATCAAGGGTGAGAT
      Db      2826 CTCTCAAGTTCATAGCTGTGTGATACAAGCTTCACTGTGACATCCAAGGGCGAGAC
      Oy      3048 GATGAGAGTTCATAGAGCTATATACATTCAAATGAGCTACCTTATTTGATACCTCCTACC
      Db      2886 GACGAGAGTTCATAGAGCTGTACACATTCAAATGAGCAACCTTCTGATACCTCCGACA
      Oy      3108 ACCTTGCTTCTATGAACTTCAATGTTGTGTGCTGCTGCGTTCAAATGCGATCAATTAAC
      Db      2946 ACCCTGCTCTACTGAACCTTCAATGAGTGTAGCTGCGATCTCAATGCGATCAACAAC
      Oy      3168 GGATATGAGTATGAGGGCCCCCTTTGGGAAGCTATTCTTTGCATTTGGGTGATTGTC
      Db      3006 GGATATGATCATGAGGGCCCCCTGTTCGGGAAGCTCTTCTTGCAATTTGGGTGATTGTC
      Oy      3228 CATCTTATCCCTTCTCAAGGTTTGTGGAAGGCAAAACAGACACCAACGATTGTC
      Db      3066 CATCTTATCCCGTCTCTCAAGGCTGTGTTGGGAGGCAAGACGCCAACGATTGTC
      Oy      3288 ATGCTGTGTTCAATTTCTGCTGCTTCAATCTTCTGCTCCTTTGGGTTGGATTGATCCT
      Db      3126 ATGCTGTGTTCAATCTCTGCTGCTTCAATCTTCTGCTCCTTTGGGTTGGATTGATCCT
      Oy      3348 TTCTTGGCAGAGATGATGTCCTGCTTCTTGAAGAGTGTGTTGATGCACTAGAT
      Db      3186 TTCTTGGCAGAGATGATGTCCTGCTTCTTGAAGAGTGTGTTGATGCACTAGAT
      Oy      3408 GTCAGTGATCAGCTCCCCC 3427
      Db      3246 GTCAAGCAGTGAATTCCCC 3265

RESULT 7
US-10-160-719-1
; Sequence 1, Application US/10160719
; Publication NO. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (63) ... (3239)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3487) ... (3487)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (3568)
; OTHER INFORMATION: n = A,T,C or G
US-10-160-719-1

Query Match      64.7%; Score 2546.8; DB 13; Length 3568;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 2829; Conservative 0; Mismatches 422; Indels 9; Gaps 2;

      Oy      177 CACAACCGCAACGAGCTGTCGTATCCCGCGCGAGCGCGATCCCGGCGGAGCGCGC
      Db      6 CCAACGCGTCCGAGCTCGTCGTATCCCGCGCGAGCGCGATCCCGGCGGAGCGCGC
      Oy      237 CCGGAGCAGAACGGGAGGTGTCCAGATTGCGGCGGAGCGAGCTCGGCTTGCCCCCGGC
      Db      66 GACCAAGCGGAACGGCCAGGTGTCCAGATTGCGGCGGAGCGAGCTCGGCGGAGCGCGC
      Oy      297 GGGGACCCCTTCTGTCGCTGCAACGAGTGGCGCTTCCCGCTGCGGAGCTGTACGAA
      Db      126 GGGGAGCTTTCTGTCGCTGCAACGAGTGGCGCTTCCCGATCTGCGGAGCTGTACGAG
      Oy      357 TAAGAGCGCCGGAGGCGCAGCAAGACTGCCCCCAAGTCAAGACTCGATACCAAGCGCTC
      Db      186 TAGAGCGCCGGAGGCGCAGCAAGACTGCCCCCAAGTCAAGACTCGATACCAAGCGCTC
      Oy      417 AAGGCTGCCAACGTGTGACCGGTGACGAGAGAGAGACGGCGTGATGACCTGACCAAC
      Db      246 AAGGGGTGCGCGCGCGTGCCTGCGGAGCAGAGAGAGAGCGGCTGACGACCTGAGCAAC
      Oy      477 GAGTTCAACTGG--GACGGCCATGACTCGCAGTCTGTGGCCGAGTCCATGCTTACGGC
      Db      306 GAGTTCAACTGGAGCGACAGCAAGACTCCCAATGACTTGGCGAGTCCATGCTTACGGC
      Oy      534 CACATGAGCTACGGCGGTGAGAGTGACCTTAATGGCGGCCCAAGCTTCCAGTCAAC
      Db      366 CACATGAGCTACGGCGCGCGCGCGACCTGACCGCGTGGCGGACCATTCACCCCATC
      Oy      594 CCGAATGTTCCACTCTCTCAACCAAGGCGCAATGTGTGATGACATCCACCGAGAGCAC
      Db      426 CCGAATGTTCCCTCTCTCAACCAAGGCGAGATGTGTGATGACATCCCGCGAGAGCAC
      Oy      654 GCGCTGTGCTCTTCTTCAATGGGTGTGGGGGAAGAGATACATCCCTTCTTATGCG
      Db      486 GCCCTTGTGCTCTCTTCTTCAATGGGTGTGGGGGAAGAGATACATCCCTTCTTATGCG
      Oy      714 GATCCCAAGCTTACCTGTGCAACCCAGTCTATGAGACCATCCAGATCTTGTGCTATAT
      Db      546 GATCCCAAGCTTCTCTGTGCAACCCAGTCTATGAGACCATCCAGATCTTGTGCTATAT
      Oy      774 GGGTATGTTAGTGTGCTTGAAGAAAGGATGAGATTGAAGCAGAGACAGAGAGG
      Db      606 GGGTATGTTAGTGTGCTTGAAGAAAGGATGAGATTGAAGCAGAGAGAGAGAGG
      Oy      834 ATGCAACGAGCGGGAATGATGTTGTTGATGATGATGATGATGATGATGATGATGATGAT
      Db      666 ATGCAACGAGCGGGAATGATGTTGTTGATGATGATGATGATGATGATGATGATGAT
      Oy      894 ATGATGAAGCAAGCAACAACCTGTCCAGGAAATTCACCTTCATCAAGCCAGATTAT
      Db      726 ATGATGAAGCTAGACAGCATTTGTCCAGAAAGATCCCGCTTCAAGCCAGATTAT
      Oy      954 CCATATAGGATGATTAATTAATTCGGCTGTGTTGGGTTCTTCTTCCACTACCGA
      Db      786 CCCTATAGGATGATTAATTAATTCGGCTAGTGTGTTGTTCTTCTTCCACTACCGA
      Oy      1014 GTGATGATCCGGTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGAAATC
      Db      1073 GTGATGATCCGGTGAATGATGATTTGCTTGTGCTCATATCTGTTATCTGAAATC
```

Db 846 GTGATGCATCCGGTGCCGTGATGCATTGCTTATGGCTCATATCTGTGATCTGTGAATT 905  
Qy 1074 TGGTTTCCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGTCCCTATTGAGAGAG 1133  
Db 906 TGGTTTCCCATGTCTTGGATTCTTGATCAATTCCCAAAGTGTGTCTATCGAGAGGGAA 965  
Qy 1134 ACTTACTAGACCCGGCTGTCACTGAGGTTCCACAAGGAAGGCCCATCTCAACTTGTCT 1193  
Db 966 ACCATCTTGAACGGCTGAGTTTAAGGTTTGACAAGGAAGGGCATCTCTCAACTCGCC 1025  
Qy 1194 CCAATTGATTTCTTGTGACGTAACGGTTGATCCCTTAAAGAACTCTTGTGACACAACA 1253  
Db 1026 CCGTTGATTTCTTGTGACGTAACGGTTGATCCCTTGAAGGAACCTCCATTGTCACGTCT 1085  
Qy 1254 AATACGTCTATCTATCTCTTCCGTGGAATTATCCTGTGATAGGTTTCTTGTCTATGTT 1313  
Db 1086 AATACGTCTATCTATCTCTTCCGTGGAATTATCCTGTGATAGGTTTCTATGCTACGTT 1145  
Qy 1314 TCTGATGATGCTGCTGCAATGCTAACGTTTGAAGCATTTATCTGAACATCTGAATTGCA 1373  
Db 1146 TCTGATGATGCTGCTGCCATGCTGACATTTGAAGCATTTGTCTGAACATCTGAATTGCA 1205  
Qy 1374 AAGAATGGGTTCTTCTGCAAAACGGTACATATTGAACCTCGCGCTCCAGAGTGTAC 1433  
Db 1206 AAGAATGGGTTCTTCTGCAAAAGATATAGCCTTGAGCCTCGTCTCCAGAGTGTAC 1265  
Qy 1434 TTCCAACAGAGATAGACTACTTGAAGACAAGGTGGCAGCAAACTTTGTTAGGAGAGG 1493  
Db 1266 TTCCAACAGAGATAGACTACTTGAAGACAAGGTGGCGCCAAACTTGTGTAGAACGG 1325  
Qy 1494 AGAGCAATGAGAGAGATGATGAGGAATTCAAGGTGAGATCAATGCTTAGTGCCAAA 1553  
Db 1326 AGAGCAATGAGAGAGATGATGAGGAATTCAAGGTGAGATCAATGCTTAGTGCTAAA 1385  
Qy 1554 GCCCAGAAAGTCTCTGAAGAAGATGACAATGCAAGTGAACCCCTGCGCTGGAAC 1613  
Db 1386 GCCCAGAAAGTCTCTGAGGAAGATGACAATGCAAGTGAACCTGCGCGGAAT 1445  
Qy 1614 AATGTTCCGTATCATCTCTGGAATGATTCAGTCTTCTTGCCAAAGCGGAGGCTTGAC 1673  
Db 1446 AATGTTCCGTATCATCTCTGGAATGATTCAGTCTTCTTGCTCAAGTGTGGCCATGAT 1505  
Qy 1674 TGTGAGGAAAATGAATCTGCCACGATTGTTTATGTTCTAGAGAGAAAAGCAGGCTAT 1733  
Db 1506 GTGAAGGAATGAGCTGCCCTGATGGTTATGTTTCAAGAGAAAAGCGCCAGGCTAC 1565  
Qy 1734 AACCATCATAGAAAAGCTGGTGTATGAATGCAATGCTGCCAGTCTGCTGTACTAACA 1793  
Db 1566 AACCATCATAGAAAAGCTGGTGTATGAATGCAATGCTGCCAGTCTGCTGTACTAAT 1625  
Qy 1794 AATGCTCCATTTTGTAACTTGATTTGATCACTACATCAACAACAGCAAGGCTATA 1853  
Db 1626 AATGCTCCATTTTGTCTGAACCTTGATTTGATCACTATATCAATAATAAGGCTATA 1685  
Qy 1854 AAGGAAGCAATGCTTTTATGATGAGACCCCTTTACTAGAAAGAGGTTTGTCTATGACAG 1913  
Db 1686 AAGGAAGCAATGCTTTTATGATGATCCCTTTGCTTGAAAGAAAGTTTGTCTATGTGACAG 1745  
Qy 1914 TTCCCTCAAGATTTGATGGATTTGATCCGATGACCGATATGCTAACCGGAATGTTGTC 1973  
Db 1746 TTTCTCAAGATTTGATGGGATTTGATCCGATGATGATATGCTAACAGAAATGTTGTC 1805  
Qy 1974 TTTTGTATATCAACATGAAGGTTTGTGATGATTCAGGGTCCAAATTTATGTTGTACT 2033  
Db 1806 TTTTGTATATCAACATGAAGGTTTGTGATGATTCAGGGTCCAAATTTATGTTGTACT 1865  
Qy 2034 GGATGTATTTGAAGGACGATTAATGTTATGATGCCCCCAAAACAAGAGCCA 2093  
Db 1866 GGATGTATTTGAAGGACGATTAATGTTATGATGCTCCCAAAACAAGAGCCA 1925  
Qy 2094 CCATCAAGGACTTGCACTGCTGGCCCAAGTGTGCTTTGCTGCTTGTGCAAT 2153  
Db 1926 CCATCAAGGACTTGCACTGCTGGCCCAAGTGTGCAATTTGCTGCTTGTGTAAC 1985

Qy 2154 AGGAAGCAAAAGAACTAACCAAAACCCAAAACAGACA-----AGAAAAAGTTATTTT 2207  
Db 1986 AGGAAGCAAAAGAAAGAACCAAGACCTCTAAACCTTAATTGAGAAAGATAAAGAACTT 2045  
Qy 2208 TTCAAGAAAAGAAAGAACCAATCCCTGATATGCTCTGTGTAATTGACGAAGCTGCT 2267  
Db 2046 TTTAAGAAAAAGAAATCAAGCCCTGATATGCTCTGTGTAATTGATGAGCCGCT 2105  
Qy 2268 CCAGAGCTGAGAAATGAAAAGCCGGTATTTGTAATCAACAAAAATTAGAAAAGAAATTT 2327  
Db 2106 CCAGAGCTGAGAAATGAAAAGCTAGTATTTGTAATCAACAGAGTTGAAAAGAAATTT 2165  
Qy 2328 GGCCAATCTTCTGTTTGTATACATCCACACTTCTCGAGAATGTGGAACCTGAGAAGT 2387  
Db 2166 GGCCAGTCTTCAAGTTTGTGATCCACACTTCTTGAAGAATGTGGAACCTGAGAAGT 2225  
Qy 2388 GCAAGTCTGCTCTCTTTGAAAAGAGCTATACATGTCATTAGTTGTGTTATGAAGAC 2447  
Db 2226 GCCAGTCCAGCTTCTCTTGAAGAGAGCTATACATGTCATTAGTTGTGATGAAGAC 2285  
Qy 2448 AAGACAGACTGGGAAAAGATTTGGCTGATCTATGATCAGTTACAGAAAGATTTCTA 2507  
Db 2286 AAAACAGGCTGGGAAAAGATTTGGTGTGATTTAGATCAGTCACAGAAAGATTTCTT 2345  
Qy 2508 ACTGTTTCAAGATGATGTCATGTTGCGGTCATTTACTGCATACCTTAACGGGTT 2567  
Db 2346 ACTGGTTTAAAGATGACATGCTAGTTGGCGGTCAATTTACTGCATACCTTAACGGGCC 2405  
Qy 2568 GCAATCAAGGTTCTGCACTCTGATCTTTCAGATGCTCTTCAACAGGTGCTCGGTGG 2627  
Db 2406 GCCTTCAAGGTTCCGCACTCTCAATCTTCCGATGCTTTTCAACAGGTTCTTCGTTGG 2465  
Qy 2628 GCTCTTGGGCTATTTGAGATCTTCTTCAAGCAATCATTTGCCCTTTGGTATGGGTATGCT 2687  
Db 2466 GCTCTTGGGCTATTTGAGATCTTCTTCAAGCAATCATTTGCCCTTTGGTATGGGTATGCT 2525  
Qy 2688 GGGGCTGAAAATTTTGAAGATTTTCTTACATCAATCTCATGCTGATCTTGAGACA 2747  
Db 2526 GGTGACTAAAGTTCCTGGAAGGTTTCTGATCATTAATCTCATGCTGATACCTTGAGACA 2585  
Qy 2748 TCTATTTCCCTCTTGGCTTACTGTACATTTGCCCTGCAATCTGTTTATGACAGGAAATTT 2807  
Db 2586 TCTATTTCCCTCTTGGCTTACTGTACATTTGCCCTGCAATCTGTTTATGACAGGAAATTT 2645  
Qy 2808 ATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGTTCAATGTCATTTTATCTGCATT 2867  
Db 2646 ATCACTCCAGAGCTGAATTAATGTTGCCAGCTGTGTTCAATGTCATTTTATCTGCATT 2705  
Qy 2868 TTTGCTACGAGCATCTTGAATGAATGAGATGAGTGTGTTGAATTTGATGACTGTGAGAG 2927  
Db 2706 TTTGCTACGAGCATCTTGAATGAATGAGATGAGTGTGTTGAATTTGATGACTGTGAGAG 2765  
Qy 2928 AATGACGAGTTCTGGGTCATTTGAAGGTGTGCTTCAACCTCTTGTGTTTCCAGGGA 2987  
Db 2766 AACGACGAGTTTGGGTCATTTGAAGGTGTGCTTCAACCTCTTGTGTTTCCAGGGA 2825  
Qy 2988 CTTCTCAAGGTCATAGCTGTGTTGATACAGCTTCAACGTCATCAAAAGGAGTGAAGT 3047  
Db 2826 CTTCTCAAGGTCATAGCTGTGTTGATACAGCTTCAACGTCATCAAAAGGAGTGAAGT 2885  
Qy 3048 GATGAGAGTTCTCAGAGCTAATACATTTCAATGAGATCACTTATGATACCTCTTACC 3107  
Db 2886 GATGAGAGTTCTCAGAGCTAATACATTTCAATGAGATCACTTATGATACCTCTTACC 2945  
Qy 3108 ACCTTCTTCTATTTGAATTTGTTGTGTGCTGCTGCTTCAATGCGATCAATAAC 3167  
Db 2946 ACCTTCTTCTATTTGAATTTGTTGTGTGCTGCTGCTTCAATGCGATCAATAAC 3005  
Qy 3168 GGATATGATCATGGGGCCCCCTTTTGGGAAGCTATTTTGGCATTTTGGGTGATGTC 3227  
Db 3006 GGATATGATCATGGGGCCCCCTTTTGGGAAGCTATTTTGGCATTTTGGGTGATGTC 3065



QY 3228 CATCTTTATCCCTTCTCAAAAGTTTGGTTGGAAAGGCAAAACAGACACCAACGATTGTC 3287  
|||||  
Db 3066 CATCTTTACCCGTTCTCTCAAGGGCTGTGGTTGGAGGCAGAACAGACGCCAACGATTGTC 3125  
  
QY 3288 ATCGTCTGGTCCATTCTGTGGCTTCAATCTTCTCGCTCCCTTGGGTTCCGATTGATCCT 3347  
|||||  
Db 3126 ATGTCTGTGTCATCCTCTGTGGCTTGCATCTTCTCGCTGCTTGGGTTCCGATCGACC 3185  
  
QY 3348 TTCCTTGGGAAGATGATGTCGCTTCTTGAGAGAGTGTGTTGGAATTGCACTAGAT 3407  
|||||  
Db 3186 TTCCTTGGGAAGATGATGTCGCTTCTTGAGAGAGTGTGTTGGAATTGCACTAGAG 3245  
  
QY 3408 GTCAGTCATCAGCTCCCC 3427  
|||||  
Db 3246 GTCAGCAGCTGACTTCCCC 3265

RESULT 8  
US-09-900-237-7  
: Sequence 7, Application US/09900237  
: Patent No. US20020120124A1  
: GENERAL INFORMATION:  
: APPLICANT: Allen, Stephen  
: TITLE OF INVENTION: Plant Cellulose Synthases  
: FILE REFERENCE: B1170 US CIP  
: CURRENT APPLICATION NUMBER: US/09/900,237  
: CURRENT FILING DATE: 2001-07-06  
: PRIOR APPLICATION NUMBER: 60/092,844  
: PRIOR FILING DATE: 1998-07-14  
: PRIOR APPLICATION NUMBER: PCT/US99/15871  
: PRIOR FILING DATE: 1999-07-13  
: PRIOR APPLICATION NUMBER: 09/720383  
: PRIOR FILING DATE: 2000-12-21  
: NUMBER OF SEQ ID NOS: 33  
: SOFTWARE: Microsoft Office 97  
: SEQ ID NO 7  
: LENGTH: 3786  
: TYPE: DNA  
: ORGANISM: Zea mays  
US-09-900-237-7

Query Match 56.9%; Score 2241; DB 10; Length 3786;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;  
  
QY 123 GGAAGTGGAGGGAGGAAGCGATGAGCGGCGGCTGGTGGCCGGCTCCCAAC 182  
|||  
Db 194 GGGCGGGGGGGGGGCTGAGATGAGGCTAGCGGGGCTGGTGGCCGGCTCCCAAC 253  
  
QY 183 CGCAACGAGCTCGTCTCATCCGCGCGACGGCGA-TCCCCGGCCGAAGCCGGCGCGGA 241  
|||  
Db 254 CGGAACGAGCTGTGTGATCCGCGCGACCGGAGTCCGGAGCCGGCGCGCGCG 313  
  
QY 242 GCAGAACGGG-AGGTGTGCAAGATTGGCGCGACGAGCTGGCCTTGCCCCGGC 296  
|||  
Db 314 GCGCGCGGGGGCGAGGCGCGCGTGCAGATATGCGGCGACGAGTCCGGGTGGCTTCGAC 373  
  
QY 297 GGGGACCCCTTCTGTGGCTGCAAGAGTGGCCTTCCCGTCTGCGGGACTGTACGAA 356  
|||||  
Db 374 GGGGAGCCCTTCTGTGGCTGCAAGAGTGGCCTTCCCGTCTGCGCGCTGTACGAG 433  
  
QY 357 TACGAGCGCGGGAGGGGACGCAAGTGCAGTCAAGTCAAGTCAAGCGCCTC 416  
|||||  
Db 434 TACGAGCGCGGGAGGGGCTCGCAAGCTGCGCAGTCAAGTCAAGCGCCTC 493  
  
QY 417 AAGGCTGCCAACGTGTGACCGGTGACGAGAGAGGACGCGCTGATGACTGACAA 476  
|||||  
Db 494 AAGGCTGCCCGCGGGTGGCGCGGACGAGAGAGGACGCGCTGACGACTGAGAGG 553  
  
QY 477 GAGTTCAACTGG-----GACGGCCATGACTCGAGTCTGTGGCCGAGTCC 521  
|||||  
Db 554 GAGTTGGGCTTGACGAGCGCGCGGCCCAAGAGACGACCCCGAGTACGTGCGCGAGTCC 613

QY 522 ATGCTTACGGCCACATGAGCTACGGCCGTGGAGTGACCTTAATGCGCGCCCAAGCT 581  
|||||  
Db 614 ATGCTACGGGCGCAGATGAGCTACGGCCGTGGAGTGACCTTAATGCGCGCCCAAGCT 664  
  
QY 582 TTCCAGCTCAACCCCAAGTTCACCTCCTCACCACGAGCAATGTGATGACATCCCA 641  
|||||  
Db 665 TTCAGCCCCGTCCCAAGTTCACCTCCTCCTCACCACGAGCAATGTGATGACATCCCG 724  
  
QY 642 CCGAGCAGCAGCGCGTGTGCTTCTTATG-----GGTGTGGGGAAGAG 692  
|||||  
Db 725 CCGAGCAGCAGCGCGTGTGCTTCTTATG-----GGTGTGGGGAAGAG 784  
  
QY 693 ATACATCCCTTCTTATGCGGATCCCAAGCTTACCTGTGCAACCCAGGTCTATGACCCA 752  
|||||  
Db 785 ATCAACCCGCTCCTTTCGAGATCCCAAGCTTCCAGTGCACCGAGATCCATGACCCG 844  
  
QY 753 TCCAAAGATCTTGTGCTGATGAGGATGATGATGATGATGATGATGATGATGATGAT 812  
|||||  
Db 845 TCCAAAGATCTTGTGCTGATGAGGATGATGATGATGATGATGATGATGATGATGAT 904  
  
QY 813 TCGAAGCAGACAAAGAGAGATGACACGAGCGGGAATGATGATGATGATGATGATGAT 869  
|||||  
Db 905 TCGAAGCAGACAAAGAGAGAGCGGCTGACGATGACGAGCGAGGAGGAGGATGATGAT 964  
  
QY 870 GGTGACGATGCTGATCTACCACTAATGATGATGATGATGATGATGATGATGATGAT 929  
|||||  
Db 965 GGTGACGATGCTGATCTACCACTAATGATGATGATGATGATGATGATGATGATGAT 1024  
  
QY 930 CCACTTCCATCAAGCCAGATTAATCCATATAGATGATTAATTAATTAATTAATTAAT 989  
|||||  
Db 1025 CCACTTCCATCAAGCCAGATTAATCCATATAGATGATTAATTAATTAATTAATTAAT 1084  
  
QY 990 TTGGGTTCTTCTTCCACTACGAGTATGATGATGATGATGATGATGATGATGATGAT 1049  
|||||  
Db 1085 TTGGGTTCTTCTTCCACTACGAGTATGATGATGATGATGATGATGATGATGATGAT 1144  
  
QY 1050 CTCATATCTGTATCTGAAATCTGTTTCCATGCTTGGATCTTGAATCAATCCCA 1109  
|||||  
Db 1145 CTCATATCTGTATCTGAAATCTGTTTCCATGCTTGGATCTTGAATCAATCCCA 1204  
  
QY 1110 AAGTGTCCCTATGAGAGAGACTTACCTAGCCGGCTGCTACGTAGGTTGCAAG 1169  
|||||  
Db 1205 AAGTGTCCCTATGAGAGAGACTTACCTAGCCGGCTGCTACGTAGGTTGCAAG 1264  
  
QY 1170 GAAGCCAGCCATCTCACTTGTCTCAATGATTTCTTGTCACTACGTTGATCCCTTA 1229  
|||||  
Db 1265 GAAGCCAGCCATCTCACTTGTCTCAATGATTTCTTGTCACTACGTTGATCCCTTA 1324  
  
QY 1230 AAGGAACCTCTTGTGCAACAATACTGTTCTATCTATCTTCTGATGATTAATCCT 1289  
|||||  
Db 1325 AAGGAACCTCTTGTGCAACAATACTGTTCTATCTATCTTCTGATGATTAATCCT 1384  
  
QY 1290 GTTGAATGAGTTCTTGTATGTTTCTGATGATGATGATGATGATGATGATGATGAT 1349  
|||||  
Db 1385 GTTGAATGAGTTCTTGTATGTTTCTGATGATGATGATGATGATGATGATGATGAT 1444  
  
QY 1350 TTATCTGAACATCTGAATTGCAAAAGAAATGGTCTTCTGCAACGCTACATATT 1409  
|||||  
Db 1445 TTATCTGAACATCTGAATTGCAAAAGAAATGGTCTTCTGCAACGCTACATATT 1504  
  
QY 1410 GAACCTCGGCTCCAGAGTGTACTTCCACAGAGATGACTACTTGAAGACAGAGT 1469  
|||||  
Db 1505 GAACCTCGGCTCCAGAGTGTACTTCCACAGAGATGACTACTTGAAGACAGAGT 1564  
  
QY 1470 GCAGCAAACTTGTAGGAGAGAGAGCAATGAAGAGAGATGAGGAATTCAAGGT 1529  
|||||  
Db 1565 GCTGCTTCAATTTGTAGGAGAGAGAGCAATGAAGAGAGATGAGGAATTCAAGGT 1624  
  
QY 1530 AGAATCAATGCTTATGTTCCAAAGCCCAAGAAAGTTCTGAAGAGATGACAAATGCA 1589  
|||||  
Db 1625 AGAATCAATGCTTATGTTCCAAAGCCCAAGAAAGTTCTGAAGAGATGACAAATGCA 1684  
  
QY 1590 GATGGAACCCCTGCTGGAACCAATGTTGTGATCATCTGGAATGATTGAGTCTTC 1649

```
Db      1685 GATGGAAAGCCCTGGCCCTGGAAACACGTACCGCATCATCTTGAATGATTCAGGTTATTC 1744
Qy      1650 CTTGGCCAAAGCGGAGGCTTGACTGTGAGGGAAATGAACCTGCCAGATTGGTTATGTT 1709
Db      1745 CTTGGCCAAAGTGGCGGTGCTGATGTGAAGAAATGAGTTGCTCGCCTGGTTTATGTC 1804
Qy      1710 TCTAGAGAAAACGACGAGCTATTAACATCATAGAAGAGCTGTGCTATGATGATG 1769
Db      1805 TCGAGAGAAAAGAGCGCAGTTATTAACATCAACAAGAGCTGTGCCATGATGACACTG 1864
Qy      1770 GTCCGAGTCTGTGCTGTACTTAACAATGCTCCATATTGTTGTTAACTTGATTGTGATCAC 1829
Db      1865 GTCCGTGTCTGTGCTGTCTTATCAAAATGCTGCATACCTATTGAAGCTTGACTGTGATCAC 1924
Qy      1830 TACATCAACAACAGCAGGCTATTAAGAACATGTGTTTATGATGAGCCCTTACTA 1889
Db      1925 TACATCAACAATAGCAAGGCCATTAAGAAGGCTATGTGTTTCAATGATGATCCTTTGGTG 1984
Qy      1890 GGAAGAAAGTTTGTCTATGTATGACAGTTCCCTCAAAAGATTGATGGGATTGATCGCCATGAC 1949
Db      1985 GGAAGAAAGTGTGTATGTATGACAGTTCCCTCAGAGTTGATGTTATGACAAAATGAT 2044
Qy      1950 CGATATGCTAACCGGAATGTGTCTTTTGTATATCAACATGAAGGTTTGATGTTAT 2009
Db      2045 CGATACGCTAACAGGAACGTTGTCTTTTGTACATCAACATGAAGGTTTGACGCTAT 2104
Qy      2010 CAGGTCCAATTTATGTGTGCTAGTGTGATTTAGAAAGCAGCATTTATATGTTAT 2069
Db      2105 CAAGAACCATTTATGTGGTACTGATGTGTTTTCAGACGGCAGCACTGTATGTTAT 2164
Qy      2070 GATGCCCCCAAAACAAGAGCCCATCAAGGACTTGCAACTGTGGCCCAAGTGTGC 2129
Db      2165 GATGCTCTTAAACGAAGAGCCCATCAAGAACTTGCAACTGTGGCCCAAGTGTGC 2224
Qy      2130 TTTTGTGTGTGCTGCTTTGGCAATAGAAACAAAAGAGACTTACCAACCCAAACAGAG 2189
Db      2225 CTCTCTGTGCTGCTGACAGCAAGAACAGAAATAAAAGAGACTTACAAACCAAGAGCGAG 2284
Qy      2190 AAGAAAAAGTTATTTTTCAGAAAGAGAGAACCAATCCCTGCATATGCTCTTGGT 2249
Db      2285 AAGAGAAAAAGTTATTTTTCAGAAAGAGAGAAACCAATCTCTGCATATGCTTTGGGT 2344
Qy      2250 GAAATTGACGAAGCTGCTCCAGAGCTGAGAAATGAAAAGCCGTTATGTAATCAACAA 2309
Db      2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAAGAGCCGGAATCGTAATCAACAG 2404
Qy      2310 AAATTAGAAAAGAAATTTGGCCAATCTCTGTTTTTTGTACATCCACTTCTCGAAGAT 2369
Db      2405 AAACTAGAGAGAAATTTGGGCAAGTCTCTGTTTTTGTCCGATCAACACTTCTTGAGAAC 2464
Qy      2370 GGTGAACCTTGAAGAGTGCAGAGTCCCTGCTCTTTTGAAGAGCTATACATGTCATT 2429
Db      2465 GGAGGGACCTTGAAGAGCGCAAGTCCAGCTTCTTCTGAAGAGCTATACATGTTATC 2524
Qy      2430 AGTTGTGTTATGAAGACAAAGACAGACTGGGGAAAAAGAGATTGGCTGATCTATGATCA 2489
Db      2525 AGCTGCGGCTACGAAGACAAAGACCGACTGGGAAAAAGAGATTGGCTGATTTACGATCG 2584
Qy      2490 GTTACAGAGATATTTCTAAGTGTTCAGAGATGTCATGTTGCGGCTCAATTTAC 2549
Db      2585 ATCACAGAGATATCTTGAAGTATTAAGATGCACTGCCATGGCTGGCGTCTATTTAC 2644
Qy      2550 TGACATACCTAAACGGGTTGCATTCAAAGGTTCTGCACCTGTGAATCTTTCAGATGCTCT 2609
Db      2645 TGACATCCGGAAGCGGCTGCATTCAAAGGTTCTGGCGCTGTGAACCTTTCGAGCGTCTT 2704
Qy      2610 CACCAAGTGTTCGGTGGGCTCTGGGCTATTTAGATCTTCTTCAGCAATCATTTGCCCT 2669
Db      2705 CACCAAGTCTTTCGCTGGGCTTGGGCTCGAATTTCTTCAGCAAGCACTGCCCA 2764
Qy      2670 CTTGGTATGGTATGTGTGGCGTCTGAATTTTGGAAAGATTTTCTACATCAACTCC 2729
```

```
Db      2765 CTTTGTACGATACGGCGGCGGCTAAAAATTCCTGAAAAGTTTCTTATATCAACTCC 2824
Qy      2730 ATCGTATCTTGGACATCTATTTCCCTCTTGCTTACTGTACATTTGCCGCACTGT 2789
Db      2825 ATCGTTATCCCTGAGCTTCATCTCTCTCTGCTTACTGTACCTTGCCTGCCATCTGC 2884
Qy      2790 TTATTGACAGGAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGTTTCATG 2849
Db      2885 CTGCTCACGGGAAGTTATACACACAGAGCTTACCAATGTCGCCAGTATCTGTTCATG 2944
Qy      2850 TCACTTTTATCTGCAATTTTGTCTACGAGCATCTAGAAATGAGATGAGTGGTGTGGA 2909
Db      2945 GCACTTTTATCTGCACTCTCCGTGACCGGCATCTGGAATGAGGTGAGTGGCTGGCC 3004
Qy      2910 ATTGATGACTGTGAGGAATGAGCAGTCTGGGTCAATTGAGGTGTCTCACAACCTC 2969
Db      3005 ATGACGACTGTGAGGAACGAGCAGTCTGGGTCAATCGAGCGTTCGCGCATCTG 3064
Qy      2970 TTTGCTGTCTCCAGGACTTCTCAAGTCAATAGCTGTGTTGATACAAGCTTCAACCGTG 3029
Db      3065 TTCGCGGTGTCCAGGCGCTGTGAAGTGTTCGCGGCATCGACAGAGCTTCAACCGTG 3124
Qy      3030 ACATCAAGGCTGAGATGATGAGAGTCTCAGAGCTATATACATTTCAATGACTACC 3089
Db      3125 ACGTGGAAGCGCGGAGCAAGAGAGTCTCGAGCTGTACACGTTCAAGTGAACCAAC 3184
Qy      3090 TTATGTATACCTCTTACCACTTGTCTTATTTGAATTCATTGTGTGTGCTGCGCTT 3149
Db      3185 CTGCTGATACCCCGACCAAGCTCTCTCTGTGAATTCATCGGGGTGTGGCGGATC 3244
Qy      3150 TCAATGCGATCAATTAACGATATGATGATGAGGCGCCCTCTTTGGGAAGCTATCTTT 3209
Db      3245 TCGAAGCGGATCAACACGGGTACGAGTGTGGGCCCCCTGTTCCGGAAGCTTCTTTC 3304
Qy      3210 GCATTTTGGGTATGTTCATCTTTATCCCTTTCATAAGGTTTGTGGAAGCAAAAC 3269
Db      3305 GCCTTGTGGGTATGTTCACCTGTACCCGTTCTCAAGGTTCTGTGGGAGGAGCAAC 3364
Qy      3270 AGGACACCAACGATTTGATCGTCTGTGCCATTTGCTGCTTCAATCTTCTGCTCTT 3329
Db      3365 AGGACGCCGAGATCGTATGCTGTGCATCTGTGGCTCGATCTTCTGCTCCTG 3424
Qy      3330 TGGGTTCCGATTGATCTCTTCTCTTCCGAAGATGATGTTCCGCTTCTTGAGAGTGTGT 3389
Db      3425 TGGTCCCGCTGACCCCGTTCTCGCCAAGAGCAACGCCCGCTCTTGAGGAGTGTGCC 3484
Qy      3390 TTGATTGCAACT 3402
Db      3485 CTGACTGCAACT 3497
```

```
RESULT 9
US-10-209-059-17
; Sequence 17, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugge, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209, 059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3813
```





Db 2225 CTCTTGTGCTGCGAGCAGAAACAAGATAAAAAAGAGACTACAAAACCAAGACGAG 2284

Qy 2190 AAGAAAAAGTATATTTTTCAGAAAGAGAACCAATCCCTGCATATGCTTGGT 2249

Db 2285 AAGAGAAAAAGTATATTTTTCAGAAAGAGAAAAACCATCTCCTGCATATGCTTGGGT 2344

Qy 2250 GAAATTGACGAAGCTGCTCCAGAGAGCTGAGATGAAAAAGCCGGTATTTGAATCAACAA 2309

Db 2345 GAAATTGATGAAGGTGCTCCAGGTGCTGATATCGAAGAGCCCGAATCGTAATCAACAG 2404

Qy 2310 AAATTAGAAAAAGAAATTTGGCCAACTCTTGTGTTTGTATACATCCACATCTCGAGAAAT 2369

Db 2405 AAAGTAGAGAAAGAAATTTGGGCAAGTCTTGTGTTTGTGTCATCAACACTTCTTGAGAAC 2464

Qy 2370 GGTGAACTTGAAGAGTGCAGAGTCCGTCTCTTTTGAAGAAGCTATACATGTCATT 2429

Db 2465 GGAGGACCTGGAAGAGCGCAAGTCCAGCTTCTCTTGAAGAGAGCTATACATGTTATC 2524

Qy 2430 AGTTGTGTTATGAAGACAAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCA 2489

Db 2525 AGCTGCGGCTACGAAGACAAGACCGACTGGGAAAAAGAGATTGGCTGATTTACGATCG 2584

Qy 2490 GTTACAGAGATATTTCTAAGTGTTCAGAGTGCATTTGTCATGTTGGCGGTCAATTTAC 2549

Db 2585 ATACAGAGAGATATCTTGAAGTATTAAGATGCACTGCCATGGCTGGCGGTCTATTTAC 2644

Qy 2550 TGCATACCTTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAATCTTTAGATCGTCTT 2609

Db 2645 TGCATCCCGAAGCGGCTGCATTCAAGGTTCTGCGCCTTGAACTTTCCGACCGTCTT 2704

Qy 2610 CACCAGTGTCTGGTGGGCTCTGGGCTATGAGATCTTCTGAGCAATCATGCCCCCT 2669

Db 2705 CACCAGTCTTGGCTGGGCTCTGGGCTCGAAATTTCTTACAGCAAGCACTGCCCA 2764

Qy 2670 CTTGTATGGGTATGGTGGCGGTGAAATTTTGAAGAAGATTTTCTACATCAACTCC 2729

Db 2765 CTTGTATCGGATACGGCGGCGGCTAAATTTCTGGAAGGTTTCTATATCAACTCC 2824

Qy 2730 ATCGTATCCTTGGACATCTATTTCCCTCTGGCTTACTGTACATTCCTGCCATCTGT 2789

Db 2825 ATCGTTATCCTCGACGTCCATTCCTCTCGGCTTACTGTACCTTGCCATCTGC 2884

Qy 2790 TTATTGACAGGAAATTTATCACTCCAGAGCTGAATATGTTGCCAGCTGTGTTCATG 2849

Db 2885 CTGCTACCGGGAAGTTTATCACACCAGAGCTTACCAATGTCCGCAATCTGTTTATG 2944

Qy 2850 TCACCTTTTATCTGCATTTTGTCTACGAGATCTGAATGAGATGAGTGTGTTGA 2909

Db 2945 GCACCTTTTATCTGCATCTCCGTGACCGGCACTCGAAATGAGTGAAGTGGCGTGC 3004

Qy 2910 ATTGATGAGTGTGAGAGAAATGAGCAGTTCTGGGTCTATGAGGTGTCTCCACACCTC 2969

Db 3005 ATGACGACTGTGGAAGGAAGCAGAGCTGTGGGTCTATCGAGGCGTTCCGCGCATCTG 3064

Qy 2970 TTTGCTGTGTTCCAGGACTTCTCAAGGTCACTAGCTGTGTTGATACAGCTTCAACCTG 3029

Db 3065 TTGCGGTGTTCAGAGGCTGCTGAAAGTGTCCCGGCATCGACAGACTTCAACCTG 3124

Qy 3030 ACATCAAAAGGTGAGATGATGAGAGTCTCAGAGCTATATACATTCAAATGGACTACC 3089

Db 3125 ACGTCAAGGCGGGGGAAGCAGAGAGTCTCGAGCTGTACACAGTTCAAGTGACACACC 3184

Qy 3090 TTATTGATACCTCTCAACCTGTCTTATGAACCTTCAATGTTGTTGCTGCTGGCGTT 3149

Db 3185 CTGCTGATACCCCGACCAACGCTCTCTGCTGAACCTTCAATCGGGGTGTCGCGGATC 3244

Qy 3150 TCAATGCGATCAATAACGAGATGATGATGAGGCGCCCTCTTTGGAAGCTATTTCTT 3209

Db 3245 TCGAAGCGATCAACAACGAGTACGAGTGTGGGCGCCCTGTTCGGGAAGCTCTTCTTC 3304

Qy 3210 GCATTTTGGGTGATGTCATCTTATCCCTTCTCAAGGTTTGTGGAAGCAAAAC 3269

Db 3305 GCCTTCTGGGTGATCGTCCACCTGTACCCGTTCTCAAGGCTGTGTTGGGAGGAGAAC 3364

Qy 3270 AGGACCAACGATTTGATCTGTGTTCCATTTCTGCTTCAATCTTCTGCTCTT 3329

Db 3365 AGGACCGCAGCATGCTCATCTGTGTTCCATCTGCTGCTGATCTTCTGCTCTG 3424

Qy 3330 TGGGTTCGATTTGATCTTCTTCTTGGCAGAGATGATGCTTCTTGAAGAGTGTGT 3389

Db 3425 TGGGTTCGCTGACACCGCTTCTTCCCAAGAGCAAGCGCCCGCTCTGAGAGTGTGC 3484

Qy 3390 TTGATTTGCAACT 3402

Db 3485 CTGACTGCAACT 3497

RESULT 10

US-10-160-719-25

Sequence 25, Application US/10160719

Publication No. US20030167528A1

GENERAL INFORMATION:

APPLICANT: Kanwarpal S. Dhugra

APPLICANT: Timothy G. Helentjaris

APPLICANT: Benjamin A. Bowen

APPLICANT: Xun Wang

TITLE OF INVENTION: Maize Cellulose Synthases and Uses

TITLE OF INVENTION: Thereof

FILE REFERENCE: 0864

CURRENT APPLICATION NUMBER: US/10/160,719

CURRENT FILING DATE: 2002-06-03

PRIOR APPLICATION NUMBER: US/09/371,383A

PRIOR FILING DATE: 1999-08-06

PRIOR APPLICATION NUMBER: US 60/096,822

PRIOR FILING DATE: 1998-08-17

NUMBER OF SEQ ID NOS: 60

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 25

LENGTH: 3813

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: CDS

LOCATION: (215) ... (3496)

US-10-160-719-25

Query Match 56.9%; Score 2241; DB 13; Length 3813;

Best Local Similarity 81.2%; Pred. No. 0;

Matches 2691; Conservative 0; Mismatches 580; Indels 42; Gaps 6;

Qy 123 GGAAGTGAAGGGAAGGAGCATGAGGCGAGCGCCGCTGTGCGGCTCCACAAC 182

Db 194 GGGCGGGGGGGGGGCTGAGATGAGGCTAGCGCGGGCTGTGCGGCTCGCATTAAC 253

Qy 183 CGCAACGAGCTGTGTCATCCGCGCGGAGCGGA-TCCGGGCGGAAGCGCGCGGA 241

Db 254 CGGAACGAGCTGTGTGATCCGCGGAGCGGAGTCGGAAGCGCGGCGCGCG 313

Qy 242 GCAGAACGGGC-----AGGTGTCCAGATTGTGCGGCGAGAGCTTGCCCCGAGC 296

Db 314 GCGGCGGGGGCGGAGGCGCGCTGCCAGATATGCGGCGAGAGTCCGGGTGCTTCGAC 373

Qy 297 GGGGACCCCTTGTGCGGTGCAAGAGTGCCTTCCCGTCTGCGGAGCTGTACGAA 356

Db 374 GGGGACCCCTTGTGCGGTGCAAGAGTGCCTTCCCGTCTGCGGAGCTGTACGAG 433

Qy 357 TACGAGCGCGGAGGCGCAGAGACTGCCCCCAAGTCAAGACTGATCAAGCGCTC 416

Db 434 TACGAGCGCGGAGGCGCTCAGAGCGTGCCCGCAAGTCAAGAGCGCTC 493

Qy 417 AAGGCTGCCAAGCTGTGACCGGTGACGAGAGAGAGCGGCTGATGACCTGACAAC 476

Db 494 AAGGCTGCCCGCGGTGCGGCGAGAGAGAGAGCGGCTGACGACCTGAGAGGCG 553

Qy 477 GAGTTCAACTGG-----GACGCGCATGACTGCGAGTCTGTGCGGAGTCC 521

Db 554 GAGTTCGGCCTGCAGAGACGGCGCCGCCACGAGAGCACCGCACTACGTCCGCGAGTCC 613  
QY 522 ATGCTCTACGGCCACATGACTACGGCCGTGAGGTGACCTTAATGGCGGCCACAAGCT 581  
Db 614 ATGCTCAGGCGCGAGATGAGCTACGGCCGCGCGCGA-----CGCGCACCCCGGC 664  
QY 582 TTCCAGCTCAACCCCAATGTTCACCTCCTCACCAACGGGCAATGTGGATGATCCCA 641  
Db 665 TTCAAGCCCCGTCCCAACGTGCCGCTCCTCACCAACGGCCAGATGGTTGATGACATCCCG 724  
QY 642 CCGGAGCAGCACGGCGCTGGTGCCTTCTTCATG-----GGTGGGGGGAAGAGG 692  
Db 725 CCGGAGCAGCACGGCGCTCGTGCCTCTTACATGAGCGCGCGCGCGGCGGCGCAAGAGG 784  
QY 693 ATACATCCCCCTTCTTAATGCGGATCCAGCTTACCTGTGCAACCCAGGTCTAATGGAACCA 752  
Db 785 ATCCACCCGCTCCCTTTCGCAATCCCAACCTTCCAGTGCAAACGAGATCCATGGAACCG 844  
QY 753 TCCAAAGATCTTGCTGCATATGGGTATGGTATGTTGCTTGGAAGGAACGAGTGAAT 812  
Db 845 TCCAAAGATCTGCGCGCTACGGAATATGGCAGCGTGGCTTGAAGAGAAATGGAGGCG 904  
QY 813 TGGAAAGCAGAGACAAGAGAGATGCACCAAGCGGGAATGATGCTGTGAT---GAT 869  
Db 905 TGGAAAGCAGAGACAGAGAGCGCTGCAGCATGTCAAGAGCGAGGGGTGGCGGTGATGGAT 964  
QY 870 GGTGACGATGCTGATCTTACCACTAATGGATGAAGCAAGACAACACTGTCCAGGAAAT 929  
Db 965 GCGCAGCATGAGATCTGCCACTAATGATGAAGCTAAGCAGCCACTTGTCCAGAAAAGTC 1024  
QY 930 CCACTTCCATCAAGCCAGATTAATCCATATAGATGATTAATCATTAATCCGCTGTGCT 989  
Db 1025 CCTATATCATCAAGCCGAATTAATCCCTACAGATGATTAATCGTTATCCGTTGTGCT 1084  
QY 990 TTGGGCTTCTTCTTCCACTACGAGTGAATGATGATCCGGTGAATGATGATTTGCTGTGG 1049  
Db 1085 TTGGGTTCTTCTTCCACTACGAGTGAATGATGATCCGGGCAAAAGATGATTTGATGTGG 1144  
QY 1050 CTCATATCTGTATCTGTGAATCTGGTTTGCCATGTCTTGGAATCTTGATCAATCCCA 1109  
Db 1145 CTCATATCTGTATCTGTGAATCTGGTTTGCGATGTCTTGATCAATGTCCCA 1204  
QY 1110 AAGTGTTCCTATTTGAGAGAGAACTTACCTAGACCGGCTGTCACTGAGGTTCCACAAG 1169  
Db 1205 AAGTGGCTTCCAATCGAGAGAGAACTTACCTGAGCCGTTTGTCACTAAAGGTTGACAAAG 1264  
QY 1170 GAAGGCCAGCCATCTCAACTTGTCTCAATTGATTTCTTGTCAGTACGGTTGATCCCTTA 1229  
Db 1265 GAAGGTCAACCCCTCTCAGCTTGCTCCAATCGACTTCTTGTCAGTACGGTTGATCCACA 1324  
QY 1230 AAGAACTCTCTTGGTGCACAACAATACTGTCTATCTATCCTTTCGTTGATTAATCCT 1289  
Db 1325 AAGGAACCTCCCTTGGTGCACAGCAACTGTCTTCCATCCTTCTGTGATTAATCCG 1384  
QY 1290 GTTGATAAGGTTTCTTGTATGTTTCTGATGATGGTGTCAATGCTAACGTTGAAGCA 1349  
Db 1385 GTTGAGAAGGTTCTCCTGCTATGTTTCTGATGATGGTGTCAATGCTTAAGTTGAAGCA 1444  
QY 1350 TTATCTGAACATCTGAATTTGCAAAAGAAATGGGTTCTTTCTGCAACGGTACAAATAT 1409  
Db 1445 TTGTCTGAACATCTGAATTTGCAAAAGAAATGGGTTCTTTCAACAAAAGTTTAATATC 1504  
QY 1410 GAACCTCGCGCTCCAGAGTGTACTTCCACAAGAATAGACTACTTGAAGACAAAGGTG 1469  
Db 1505 GAGCTCGTGTCTCTGAGTGTACTTCCACAAGAAATAGACTACTTGAAGACAAAGGT 1564  
QY 1470 GCAGCAAACTTTGTTAGGAGAGAGCAATGAAGAGAGATGAGGAATTCAGAGGTG 1529  
Db 1565 GCTGCTTCATTTGTTAGGAGAGAGGCGGATGAAGAGAAATACGAGGAATTCAGAGTA 1624  
QY 1530 AGAATCAATGCCTTAGTTGCAAAAGCCAGAAAGTTCTGAAGAAGATGACAATGCAA 1589  
Db 1625 AGGATCAATGCCTTAGTTGCAAAAGCCCAAAAGGTTCTGAAGAAAGATGACAATGCAA 1684

QY 1590 GATGGAACCCCTGCGCTGGAAAACAATGTCGTGATCATCTCGAATGATTCAGGCTTC 1649  
Db 1685 GATGGAAGCCCTGCGCTGGAAAACAATGTCGCGATCATCTCGAATGATTCAGGATTC 1744  
QY 1650 CTTGGCCAAAGCGGAGGCTTGACTGTGAGGGAATGAACCTGCCACGATGGTTATGTT 1709  
Db 1745 CTTGGCCAAAGTGGCGGTGTGATGTGAGGAATGAGTTGCCCTGCGGTGTATGTC 1804  
QY 1710 TCTAGAGAAAAGCAACGAGCTATTAACCATTAAGAAAAGCTGGTGTATGAATGCATTG 1769  
Db 1805 TCGAGAGAAAAGAGCCAGGTTATTAACCATCAAGAAGGCTGGTGCATGAATGCACTG 1864  
QY 1770 GTCCGAGTCTGCTGTACTAACAATGCTCCATATTGTTAACTTGAATGTGATCAC 1829  
Db 1865 GTCCGTCTCTGCTGTCTTATCAATGCTGCATTAACCTTGAACCTGTGATCAC 1924  
QY 1830 TACATCAACAACAGCAAGGCTATTAAGGAACAATGTGTTTATGATGACCCCTTACTA 1889  
Db 1925 TACATCAACAATAGCAAGGCCATTAAGAGGCTATGTGTTTATGATGATGATCCTTGGTG 1984  
QY 1890 GGAAGAAGGTTGCTATGTACAGTTCCTCAAAAGATTGATGGATGATCGCCATGAC 1949  
Db 1985 GGAAGAAGGTTGCTATGTACAGTTCCTCAAGAGTTGATGTATGACAAAATGAT 2044  
QY 1950 CGATATGCTAACCGGAATGTTGCTTTTGTATATCAACATGAAGGTTTGATGTAT 2009  
Db 2045 CGATAGCTTAACAGGAACGTTGCTTTTGTATCAATCAACATGAAGGTTTGAACGATAT 2104  
QY 2010 CAGGTCCAATTAATGTTGGTACTGATGTATTTAGAAGCGAGCATTAATGCTTAT 2069  
Db 2105 CAAGAGCCCATTTATGTGGGTACTGATGTGTTTTCAGACGGCAGCACTGTATGCTTAT 2164  
QY 2070 GATCCCCCAAAACAAGAACCAACCATCAAGAACTTGCAACTGCTGGCCCAAGTGTGC 2129  
Db 2165 GATGCTCCTAAACGAAGAACCAACCATCAAGAACTTGCAACTGCTGGCCCAAGTGTGC 2224  
QY 2130 TTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAAGTACCAAAACCAAAACAGAG 2189  
Db 2225 CTCTCTGTGCTGCAGCAAGAACGAATTAAGAAAGAAAGTACCAAAACCAAAACAGAG 2284  
QY 2190 AAGAAAAAGTTATTTTTCAGAAAGAGAAACCAATCCCTGCAATATGCTCTTGGT 2249  
Db 2285 AAGAAAGAAAGTATTTTTCAGAAAGAGAAACCAATCCCTGCAATATGCTTGGT 2344  
QY 2250 GAAATTGACGAAGCTCTCCAGAGCTGAGATGAAGAAAGCCGATTTGTAATCAACAA 2309  
Db 2345 GAAATTGATGAAGTCTCCAGGTCTGATATCGAAGAGCCGGAATCGTAATCAACAG 2404  
QY 2310 AAATTGAAAAAGAAATTTGGCCAATCTTGTGTTTGTATCATCCACTTCTCGAAGAT 2369  
Db 2405 AAATAGAGAGAAATTTGGGCAATCTTGTGTTTGTGCGATCAACACTTCTTGAGAAC 2464  
QY 2370 GGTGAACCTTGAAGAGTGCAGATCCTGCTTCTTTTGAAGAAAGTATACATGTCAAT 2429  
Db 2465 GGAGGAGCCCTGAAGAGCGCAAGTCCAGCTTCTTCTGAAGAAAGTATACATGTATC 2524  
QY 2430 AGTTGTGTTATGAGACAAGACAGACTGGGGAAGAAAGATTTGCTGATCTATGATCA 2489  
Db 2525 AGCTGCGGCTACGAAGACAAGACGACTGGGGAAGAAAGATTTGCTGATCTATCGATCG 2584  
QY 2490 GTTACAGAGATATTTAACTGTTTCAAGATGATTTGTCAATGTTGGCGGTCAATTTAC 2549  
Db 2585 ATCAGAGAGATATCTGAAGTATTAAGATGCACTGCCATGGCTGGCGGTATTTAC 2644  
QY 2550 TGCATACCTAAACGGGTGCATTCAAAGGTTCTGCACTCTGAATCTTTAGATCGTCTT 2609  
Db 2645 TGCATCCCGAAGCGGCTGCATTCAAAGGTTCTGCGCTCTGAACCTTTCCGACCGTCTT 2704  
QY 2610 CACCAGGTGCTTGGTGGGCTCTTGGGTCTAATGAGATCTTCTTACGAATCATTGCCCT 2669  
Db 2705 CACCAGGTCTTGGTGGGCTTGGGTCTGAAATTTTCTTCAAGAAAGCACTGCCCA 2764

QY	2670	CTTGGTATGCGGTATGTCGCCGCTCTGA AATT TTTTGAAAGATTTTCTACATCAACTCC	2729
Db	2765	CTTGGTACGGATACGCGCGCGGCTA AAAATTCCTGAAAGTTTTCTTATATCAACTCC	2824
QY	2730	ATCGTATCCTTGACATCTATTCCTCTCTGCTTACTACTACATTTGCCATCGT	2789
Db	2825	ATCGTTATCCCTGACGCTCCACTTCCTCTCTGCTTACTACTACCTTGCCATCGC	2884
QY	2790	TTATTGACAGGAAATTTATCATCTCCAGAGCTGAATAATGTTGCCAGCCTGTGTTCATG	2849
Db	2885	CTGCTCACGGGGAAGTTTATCATCACCAAGAGCTTACC AATGTCGCCAGTATCTGGTT CATG	2944
QY	2850	TCACTTTTATCTGCATTTTGTCTACGAGCATCCTAGAAA TGAGATGAGTGTGTTGGA	2909
Db	2945	GCACTTTTCATCTGCATCTCCGTGACCGGCATCCTGGAATGAGGTGAGTGGCGTGCC	3004
QY	2910	ATTGATGACTGGTGAGGAATGAGCAGTCTGGGTCATTGSAAGGTGTCTCTCACACTC	2969
Db	3005	ATCGACGACTGGTGAGGAGAACAGCAGTTCTGGGTCACTGAGGCGTTTCGGCGCATCTG	3064
QY	2970	TTTGCTGTGTTCCAGGGA CTTCTCAAGSTCATAGCTGCTGTTGATACAAGCTTCAACG TG	3029
Db	3065	TTGCGGTGTTCCAGGCGCTGCTGAAGGTGTTGCGCGGCATCGACAGAGCTTCAACG TG	3124
QY	3030	ACATCAAAGG GTGAGATGATGAGAGAGTTCTCAGAGCTATATACATTC AATGAGCTACC	3089
Db	3125	ACGTGGAAGCGCGGGGACGACGACGAGAGTTCTCGAGCTGTACACGTTCAAGTGGACCA CC	3184
QY	3090	TTATTGATACCTCCCTACCACTTGTCTTATTGAACTTCA TTGGTGTGCTGCGTGGCGTT	3149
Db	3185	CTGCTGATACCCCCGACCA CGCTCTCTCTGCTGA ACTTCA TGGGGGTGTGGCCGGGATC	3244
QY	3150	TCAATGCGATCAATTAACG GATATGAGTCATGSGGCCCTCTTTGGGAAGCTATTC TT	3209
Db	3245	TCGAACGCGATCAACAACGGGTACGAGTCGTGGGCGCCCTCTGTTCGGGAAGCTTCTT TC	3304
QY	3210	GCA TTTGGGTGATGTGCCATCTTATCCCTTCTCAAGGTTTGGTGAAGGC AAAC	3269
Db	3305	GCCTTCTGGGTGATCGTCACCTGTACCCGTTCTCCTCAAGGCTGTGTTGGGAGGCAGA AC	3364
QY	3270	AGGACACCAACGATTTGTCATCGTCTGGTCCATTCTGTGCTTCAATCTTCTGCTCCTT	3329
Db	3365	AGGACGCCGACGATCGTCATCGTCTGGTCCATCTGCTGCTCGATCTTCTCGCTCTG	3424
QY	3330	TGGGTTCCGATTGATCTTCTTCTTGCGAAGGATGATGTTCCGCTTCTTGAGGAGTGTGT	3389
Db	3425	TGGGTCCGCGTCGACCCGTTCTCTGCCAAGACGAACGCGCCGCTCTTGAGGAGTGTG GC	3484
QY	3390	TTGGA TTGCAACT 3402	
Db	3485	CTGGA CTGCAACT 3497	

```

; RESULT 11
; US-10-160-719-45
; Sequence 45, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Matize Cellulose Synth
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 45
; LENGTH: 3813
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)...(3496)
;
US-10-160-719-45

```

Query Match	56.9%	Score 2241;	DB 13;	Length 3813;
Best Local Similarity	81.2%;	Pred. No. 0;		
Matches 261;	Conservative	0;	Mismatches 580;	Indels 42;
				Gaps 6;

QY	123	GGAAGTGGAGGGGAGGAAGCCGATGGAGGCGGACCGCCGGCTGGTGGCCGGCTCCACAAAC	182
Db	194	GGGCGGGGCGGGGGGCTTGAGATGAGAGGCTAGCGCGGGCTGGTGGCCGGCTCGCATAAC	253
QY	183	CGCAACGAGCTCTGTGTCATCCGCCGCGACGGCGA-TCCGGGGCGAAGCCGGCCGGGA	241
Db	254	CGAAACGAGCTGTGTGATCCGCCGCGACCGCAGTCCGGAGCCGCGGGCGGGCGCG	313
QY	242	GCAGAACGGGC-----AGGTGTGCCAGATTGCGGCGACGAGCTGGCCTTGCCCCGGC	296
Db	314	GCGCGCCCGGCGGAGGCGCGCGTGCAGATATGCGGCGACGAGSTCGGGGTGGGCTTCGAC	373
QY	297	GGGGACCCCTTCGTGGCGTCCACGAGTGCCTTCCCGTCTCGCCGGACTGCTACGA	356
Db	374	GGGGAGCCCTTCGTGGCGTCCACGAGTGCCTTCCCGTCTCGCCGCGCTGCTACGAG	433
QY	357	TACGAGCGCCGGGAGGGGCGACGAACTGCCCCCAGTGCAGAAGCTGATACAAAGCGCTC	416
Db	434	TACGAGCGCCGGGAGGCTCCGAACGCTGCCCCGAGTGCAGAACCCGCTACAAAGCGCTC	493
QY	417	AAGGCTGCCAAACGTGTGACCGGTGACGAGGAGGAGGACGGCGTGCATGACCTGGAAC	476
Db	494	AAGGCTGCCCGCGGGTGGCCGGCGACGAGGAGGAGGACGGCGTGCAGCACTGAGGGC	553
QY	477	GAGTCAACTGG-----GACGGCCATGACTCGCAGTCTGTGCGCGAGTCC	521
Db	554	GAGTTGGCGCTGCAGAGACGGCGCCGCCACGAGGACGACCCGCAAGTACGTGCGCGAGTCC	613
QY	522	ATGCTCTACGSCCAATAGCTACGCGCGTGAAGTGAACCTAATGCGCGCCACAACT	581
Db	614	ATGCTCAGGGCGCGAGTAGCTACGGCGCGCGCGCGA-----CGCGCACCCCGGC	664
QY	582	TTCCAGCTCAACCCCAATGTTCCACTCTCTCAACCAACGGGCAATGCTGATGACATCCA	641
Db	665	TTCAAGCCCGTCCCAACGTGCCGCTCTCAACCAACGGGCAATGTTGATGACATCCCG	724
QY	642	CCGAGCAGCAGCAGCGCGTGTGCTCTTCTCATG-----GGTGTGGGGGAAGAGG	692
Db	725	CCGAGCAGCAGCAGCGCGTGTGCTCTCTCATGAGCGCGCGCGCGCGGCGGCAAGAGG	784
QY	693	ATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGTCTATGACCCA	752
Db	785	ATCCACCCGCTCCCTTTCGAGATCCCAACCTTCCAGTGAACCGAGATCCATGACCCG	844
QY	753	TCCAAGGATCTTGCTGCATATGGGTATGTAGTGTGCTTGAAGAAGACGATGAGAAT	812
Db	845	TCCAAGGATCTTGCGCGCTTACGAGATATGCGACGCTGGCTTGAAGAAGAAATGAGGGC	904
QY	813	TGGAAGCAGAGACAAAGAGAGATGCACCAAGACGGGGAATGATGCTGTGTGAT--GAT	869
Db	905	TGGAAGCAGAGACAGAGCGGCTGCAGCATGTCAAGACGAGGGGTGGCGGTATTGGAT	964
QY	870	GGTGAAGATGCTGATCTTACCACTAATGATGAAGCAAGCAACACTGTCCAGGAAAT	929
Db	965	GCGCAGATGAGATCTGCCACTAATGATGAAGCTAGGCAACCATTTGCCAGAAAAGTTC	1024
QY	930	CCACTTCCATCAAGCCAGATTAAATCCATATAGGATGATTAATTCATTTGGCTTGTGTT	989
Db	1025	CCTATATCATCAAGCCGAATTAATCCCTACAGATGATTAATGTTATCCGGTGTGTT	1084



QY	990	TTGGGGTCTCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTGTGCTTGTGG	1049
Db	1085	TTGGGTTCTTCTTCCACTACCGAGTGATGCATCCGGCGAAAAGATGCATTGTGTGG	1144
QY	1050	CTCATATCTGTTATCTGTGAAATCTGGTTTGCCATGCTTGGATTCTTGATCAATTCCCA	1109
Db	1145	CTCATATCTGTAATCTGTGAAATCTGTTTGCGATGTCCTGGATTCTTGATCAGTTCCCA	1204
QY	1110	AAGTGTTCCCTATTGAGAGAGAGACTTACCTAGACCCGGCTGTCACTGAGGTTGCACAAG	1169
Db	1205	AAGTGGCTTCCAATCGAGAGAGAGACTTACCTGAGCCGTTTGTCACTAAGGTTGCACAAG	1264
QY	1170	GAAAGCCAGCCATCTCAACTTGCTCCAATTGATTCTTTGTCACTACGGTTGATCCCTTA	1229
Db	1265	GAAAGTCAACCCCTCAGCTTGCTCCAATCGACTTCTTTGTCACTACGGTTGATCCACA	1324
QY	1230	AAGAACTCCTTTGGTCAACAATACTGTTCTATCTATCCTTTCGGTGAATTATCCT	1289
Db	1325	AAGAACTCCTCCTTGGTCAACGGAACACTGTCCTTCCATCCTTCTGTGATATCCG	1384
QY	1290	GTTGATTAAGGTTCTTGTCTATGTTTCTGATGATGGTGCTCAATGCTAACGTTGAACA	1349
Db	1385	GTTGAGAAGGTCCTCCTGTAATGTTTCTGATGATGGTGCTCAATGCTTACGTTGAACA	1444
QY	1350	TTATCTGAAACATCTGAATTTGCCAAGAAATGGTTCCTTCTGCAAACGGTACATAATT	1409
Db	1445	TTGTCTGAAACATCTGAATTTGCCAAGAAATGGTTCCTTCAACAAAAGTTAATATC	1504
QY	1410	GAACCTCGCGCTCCAGAGTGTACTTCCAACAGAAAGATAGACTACTTGAAGACAAAGTG	1469
Db	1505	GAGCCTCGTCTCCTGAGTGTACTTCCAACAGAAAGATAGACTACTTGAAGACAAAGTT	1564
QY	1470	GCAGCAACTTGTTAGGAGAGAGAGCAATGAAGAGAGATAGGAATTCAGGTG	1529
Db	1565	GCTGCTTCATTTGTTAGGAGAGAGAGCGCATGAAGAGAGAAATACGAGGAATTCAGGTA	1624
QY	1530	AGAATCAATGCCCTTAGTGCCAAAGCCCAAGAAAGTTCCTGAAGAGAGATGACAAATGCAA	1589
Db	1625	AGATCAATGCCCTTGTTGCCAAAGCCCAAAAGGTTCTGAGGAAGATGACAAATGCAA	1684
QY	1590	GATGGAACCCCTGGCCCTGAAAACAATGTTGCTGATCATCCTGGAATGATTCAGGTCTTC	1649
Db	1685	GATGGAAGCCCTGGCCCTGAAAACAACGTAACGATCATCCTGGAATGATTCAGGTATTC	1744
QY	1650	CTTGCCCAAAGCCGAGCCCTTGACTGTGAGGGAATGAATGCCACGATTGCTTATGTT	1709
Db	1745	CTTGCCCAAAGTGCGGTCGTGATGTGGAAGGAATGATGCTCGCCTGCTTATGTC	1804
QY	1710	TCTAGAGAGAAAAGCAGCGCTATAACCATCATAGAAGACTGTGCTATGATGCAATTG	1769
Db	1805	TCGAGAGAAAAGAGCCAGGTTATAACCATCACAAAGAGCTGTGCCATGAATGCACTG	1864
QY	1770	GTCCGAGTCTCTGCTGTACTAACAAATGCTCATATTTGTTAACTTGATTGATCAAC	1829
Db	1865	GTCCGTGCTCTGCTGCTTATCAATGCTGCATACCTATGAACTTGACTGTGATCAC	1924
QY	1830	TACATCAACACAGCAAGGCTATAAAGGAAGCAATGTGTTTATGATGACCCTTTACTA	1889
Db	1925	TACATCAACATAGCAAGGCCATAAAAAGGCTATGTGTTTCATGATGATCCTTTGGTG	1984
QY	1890	GGAAGAAGGTTTGCTATGTACAGTTCCTCAAGATTTGATGGGATTGATCGCATGAC	1949
Db	1985	GGAAGAAGGTGTGCTATGTACAGTTCCTCAGAGTTGATGATGACAAAATGAT	2044
QY	1950	CGATATGCTAACCGGAATGTTGCTTTTGTGATATCAACATGAAGGTTTGATGATGAT	2009
Db	2045	CGATACGCTAACAGAACGTGTCTTTTGTGACATCAATGAAGGTTTGACGGTATTT	2104
QY	2010	CAGGCTCCAATTATGTGTGTACTGATGTGTATTTGAAGCGCAGCATTATATGTTAT	2069
Db	2105	CAAGGACCCATTTATGTGTGTACTGATGTGTTTTCAGACGGCAGCATGTATGTTAT	2164
QY	2070	GATGCCCCCAAAAGAACCAACCATCAAGGACTTGCAACTGTGCGCCCAAGTGGTGC	2129

Db	2165	GATGCTCCTAAACGAAGAGCCACCATCAAGAACTTGCAACTGTGGCCCAAGTGTGC	2224
Qy	2130	TTTTGCTGTGCTGCTTTGGCAATAGGAAGCAAAAGAACTACCAAAACCAAGACAG	2189
Db	2225	CTCTCTTGTGCTGCTGCAGCAGGACAGAAATAAAAAGAACTACAACCAAGACGAG	2284
Qy	2190	AAGAAAAAGTTATTATTTTCAAGAAAGAAAGAACTAACCTCCCTGCATATGCTCTTGGT	2249
Db	2285	AAGAAGAAAAAGATTATTTTCAAGAAAGCAAGAAACCATCTCCTGCATATGCTTTGGGT	2344
Qy	2250	GAATTTGACGAAGCTGCTCCAGAGAGCTGAGAAATGAAAAAGCCGCTATTGTAAATCAACAA	2309
Db	2345	GAATTTGATGAAGGTGCTCCAGGTGCTGATATCGAAGAGCCGGAATCGTAATCAACAG	2404
Qy	2310	AAATTGAAAAAGAAATTTGGCCAACTCTCTGTTTTTGTTCATCCACACTTCTCGAAGAT	2369
Db	2405	AAACTAGAGAAAGAAATTTGGCAGTCTTCTGTTTTTGTGCGATCAACACTTCTTGAGAAC	2464
Qy	2370	GGTGAACCTTGAAAGAGTGAAGTCTGCTTCTCTTTTGAAGAAAGCTATACATGTCATT	2429
Db	2465	GGAGGACCCCTGAAGAGCGCAAGTCCAGCTTCTCTCTGAAGAAAGCTATACATGTTATC	2524
Qy	2430	AGTTGTGTTATGAAGACAGACAGACTGGGAAAAAGATTGGCTGAATCTATGATCA	2489
Db	2525	AGCTGGGCTACGAGACAGACAGACCGACTGGGAAAAAGAGATTGGCTGAATTTACGATCG	2584
Qy	2490	GTTACAGAAAGATATTCTTAACCTGCTTCAAGATGCAATGTCATGCTTGGCGGCTCAATTAC	2549
Db	2585	ATCACAGAGGATATCTTGACTGGAATTAAGATGCACTGCCATGGCTGGCGGCTATTTAC	2644
Qy	2550	TGCATACCTAAACGGGTTGCATCAAAAGTTCTGCACTCTGAATCTTTCAGATCGTCTT	2609
Db	2645	TGCATCCCGAAAGCGGCTGCATTCAAAGTTCTGCGCTCTGAACCTTCCGACCGTCTT	2704
Qy	2610	CACCAAGTCTTCGGTGGGCTCTTGGTCTATGAGATCTTCTTCAGCAATCATTTGCCCT	2669
Db	2705	CACCAAGTCTTCGGTGGGCTCTTGGTCCGTCGAAATTTCTTCAGCAAGCACTGCCCA	2764
Qy	2670	CTTTGATATGGGTATGTTGGTGGCTGAAATTTTGGAAAGATTTCTTCAATCATCAACTCC	2729
Db	2765	CTTTGATATGGGTATGTTGGTGGCTGAAATTTCTTGAAGATTTCTTATATCAACTCC	2824
Qy	2730	ATCGTATCTTGGACATCTATCCCTCTTGGCTTACTGTACATTTGCCATCTGCTGT	2789
Db	2825	ATCGTATCTTGGACATCTATCCCTCTTGGCTTACTGTACATTTGCCATCTGCTGT	2884
Qy	2790	TTATTGACAGGGAATTTTATCACTCCAGAGCTGAATATGTTGCCAGCCGTGTGTTCATG	2849
Db	2885	CTGCTACCGGGGAAGTTTATCACACAGAGCTTACCAATGTCCGCAATCTGTGTTCATG	2944
Qy	2850	TCACITTTTATCTGCATTTTGTCTACGAGCATCTTGAATGAGATGAGTGTGTGGA	2909
Db	2945	GCACITTTTATCTGCATCTCCGTGACCGGCATCTGAAATGAGTGAAGTGGCGTGCC	3004
Qy	2910	ATTGATGACTGTGAGAGATGAGCAGTTCTGGGTCAATTGAGGTGTGCTTCAACCTC	2969
Db	3005	ATGACGACTGTGTGAGAGACGACAGTTCTGGGTCAATCGAGGCGTTCCGCGCATCTG	3064
Qy	2970	TTTGTGTTGTTCCAGGGACTTCTCAAGTCAATAGCTGTGTGATACAGCTTCACCGTG	3029
Db	3065	TTTGTGTTGTTCCAGGGCTGTGAAAGTGTTCGCGCGCATCGACACGAGCTTCACCGTG	3124
Qy	3030	ACATCAAAAGGTTGAGATGATGAGAGTTCTCAGACTATATACATTCAAATGACTAC	3089
Db	3125	ACGTGAAAGGCTGGGAGCAGCAGAGAGTTCTCGAGCTGTACACGTTCAAAGTGACCA	3184
Qy	3090	TTATTGATACCTCTTACCACTTGTCTTATTGAATTCATTTGTTGTGTGCTGGCGTT	3149
Db	3185	CTGCTGATACCCCGACCAAGCTCTCTGCTGAATTCATTCGGGTGTGCGCGGATC	3244
Qy	3150	TCAAATGCGATCAATAACGATATGATCATGGGCGCCCTCTTGGAAAGCTATTTCTTT	3209

Db 3245 TCGAACGGCATCAACACGGGTACGATCGTGGGCCCCCTGTTCGGGAAGCTTTCTTC 3304  
Qy 3210 GCATTTTGGGTGATTTGCCATCTTTATCCCTTTCTCAAGGTTTGGTTGGAAGCAAAAC 3269  
Db 3305 GCCTTCTGGGTGATCGTCCACCCTGTATCCCGTTCTCAAGGGTCTGGTGGGAGGCAAAAC 3364  
Qy 3270 AGGACACCAACGATTTGTATCGTCTGTGTCATTTCTGCTGGCTTCAATCTTCTGCTCCTT 3329  
Db 3365 AGGACCGGACGATCGTATCGTCTGTGTCATCTGCTGGCTTCAATCTTCTGCTCCTG 3424  
Qy 3330 TGGGTTCCGATTTGATCTTCTTCTTGGCGAAGATGATGTCGCTTCTTGAGAGTGTGT 3389  
Db 3425 TGGTCCGCGTCGACCCGTTCTCTGCGCAAGAACACGCGCTCTGAGAGTGTGGC 3484  
Qy 3390 TTGATTTGCAACT 3402  
Db 3485 CTGACTGCAACT 3497

RESULT 12

US-09-938-842A-2530  
; Sequence 2530, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938, 842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227, 866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264, 647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300, 111  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2530  
; LENGTH: 3255  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2530

Query Match 38.3%; Score 1509; DB 10; Length 3255;  
Best Local Similarity 67.6%; Pred. No. 0;  
Matches 2207; Conservative 0; Mismatches 1030; Indels 30; Gaps 5;

Qy 144 ATGAGGCGAGCGCGGGCTGTGCGCGCTCCCAACCGCAAGCTCGTGCATC 203  
Db 1 ATGAACACCGGTGTCGTTAATCGCCGTTCTCAACAGAAATGTTGCTCAT 60  
Qy 204 CGCCCGACGCGCATCCCGGGCGAAGCCCGCGGAGAGAGCAAGCGGAGGTGCCAG 263  
Db 61 AATGCCGATGAGATGCCCCGAATAAGATCAGTCCAAGAGCTGAGTGAACATGTCAA 120  
Qy 264 ATTGCGGCGAGCGTGGCTTGGCCCCCGGGGAGACCCCTTCTGCGCTGCAACGAG 323  
Db 121 ATCTGAGAGATGAGATCGAATTGACTGTTGATGGAAGAACCGTTTGTGCATGTAAAGAA 180  
Qy 324 TGGGCTTCCCGTCTGCGCGGAGCTGACGATACGAGCGCGGAGGAGCGCAGAAC 383  
Db 181 TGTGATTTCCCTGTGTAGACCTTGTATGAGTACGAAAGAGAGGCAATCAAGCT 240  
Qy 384 TGCCCCAGTGAAGACTCGATACAGCGCTCAAGGGCTGCCAAGTGTGACCGGTGAC 443  
Db 241 TGTCCACAGTGAAGAACCCGTTTCAAAAGCTTTAAAGAGTCCAAGGTTGAAGGTGAT 300  
Qy 444 GAGAGAGAGAGCGGTGATGACCTGAGCAAGAGTCAACTGGAGGCGCATGACTCG 503  
Db 301 GAAGAGAGAGATGATGATGATTTAGACAAATGAGTTGA--GTATGAAATAATAGG 357

Qy 504 CAGTCTGCGCCGATTCATGCTCTACGGCCACATGAGCTACGGCCGTGAGTGACCT 563  
Db 358 ATTGATTTGATCAGGTTTCTGAAGGATGTCAATCTCTGTCGCAACTCCGGTTTCCA 417  
Qy 564 AATGCGCGCCACAAGCTTTCAGCTCAACCCCAATGTTCCACTCCTCACCAAGGCA 623  
Db 418 CAATCTGATTTGATTCAGCTCCACCTGCTCGATTCATTCATGCTGACTTACGGCAG 477  
Qy 624 ATGTGATGACATCCACCGAGCAGCAGCGC---TGGTCCCTTCTTCAATGGGTGT 680  
Db 478 GAGGACGTTGAGATTTCTGTATAGACATGCTTTATGTCTCTCCTTCACTTGTGT 537  
Qy 681 GGGGGAAGAGATACATCCCTTCTCTTATGCGGATCCCAAGCTTACCTGTGCAACCCAG 740  
Db 538 CATGCAATAGAGTCAATCCTGTTTCTCTTCTGACCCGACCGGTGCTGCACATCCAAG 597  
Qy 741 TCTATGACCCATCCAGAGATCTTGCTGATATGGGTATGTAATGTTGCTTGAAGAA 800  
Db 598 CTTATGTTACCTCAGAAAGATCTTGGGTTTATGTTATGGAAGTGTGCTTGAAGAA 657  
Qy 801 CGATGGAATTTGAAGCAGAGACAAGAGAGATGCAACGAGGGA-----AT 851  
Db 658 CGATGGAAGATGGAAGAGAAAGCAGAAATGAGAACTTCAGGTTGTTAGCATGAAGAA 717  
Qy 852 GATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911  
Db 718 GATCCTGATTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
Qy 912 CAATGTTCCAGAAATTCACCTTCCATCAAGCCAGATTAATCATATAGATGATATC 971  
Db 778 CCATGTTCTAGGAAGATACCAATCAATCGAGCAAGATAATCCTTACCGGATTTAAT 837  
Qy 972 ATTATTCGCTGTGTTTGGGTTCTTCTTCCACTACCGAGTGTGATCCGGTGAAT 1031  
Db 838 GTGCTACGCTTGTGATTTCTGTTCTCTTCTTCACTACCGATTTCTTCAACCCGTCAAA 897  
Qy 1032 GATGATTTGCTTGTGCTCATATCTGTTATCTGGAATCTGTTGCCATGCTTGG 1091  
Db 898 GATGATATGCTTGTGCTTATTTCTGTTATGATGATGATGATGATGATGATGATGAT 957  
Qy 1092 ATTCTGATCAATTCACCAAGTGTCTCCATTTGAGAGAGACTTACCTAGACCGCTG 1151  
Db 958 GTTCTGATCAGTTCCTTAATGTTGATCCCTATCGAGCGAAGACGTAATGACCGACTC 1017  
Qy 1152 TCACTGAGCTTGCAGCAAGAGCCAGCCATCTCAACTTGCCTCAATGATTTCTTGT 1211  
Db 1018 TCAATTAAGATATGAGAAAGAGGAAACCGTCGGAGTATCCCTGTGATGATTTGTT 1077  
Qy 1212 AGTACGCTGATCCCTTAAGAGAACCTCTTGTGTCACAAACAATACTGTCTATATC 1271  
Db 1078 AGTACAGTGCATTCATTTGAAGAGCCTCCGCTTATTTACTGCAAAATCTGTCTGTAT 1137  
Qy 1272 CTTTGGGTGATTTATCTGTTGATTAAGTTCTTGTCTATGTTTCTGATGATGTTGCTCA 1331  
Db 1138 CTTGCTGTTGATTTATCTGTGATTAAGTTGCTTGTGATGATGATGATGATGATGAT 1197  
Qy 1332 ATGCTAACGTTTGAAGCATTAATCTGAACATCTGAATTTGCAAAAGAAATGGTTCTTTC 1391  
Db 1198 ATGCTTAATTTGAAAGCTTTTCTGAGACCGCTGAATTCGCAAGAAATGGTTCTTTC 1257  
Qy 1392 TGCAACGCTACATATTAAGCTTCCGCTCCAGAGTGTACTTCCAAAGAGATGAGC 1451  
Db 1258 TGCAAGAAATATTGATTTGAGCTGCTGCCGAATGGTATTTCTGCCATAAAATGAGC 1317  
Qy 1452 TACTGAAGAAGCAAGTGGAGCAAACTTGTTAAGGAGAGAGCAATGAAGAGAG 1511  
Db 1318 TACTGAAGAATAAAGTTCATCCCGCATTTGTTAGGAGCGGCGAGCATGAAGAGAT 1377  
Qy 1512 TATGAGAAATTCAGGTGAGATCAATGCTTGTGCTTGAAGCCAGAAAGTTCTGAA 1571  
Db 1378 TATGAAGAAATTCAGAAAGTAAAGATCAATGCTTGTAGTGAACAGACAGAAAGTGCCTGAG 1437

QY	1572	GAAGGATGGACAATGCAGATGGAACCCCTGGCTTGAGAAACAATGTTCTGATCATCCT	1631
Db	1438	GATGGTTGACTATGCAGACGGTACACCTTGGCCCGTAATAGTGTGCGAGATCATCCT	1497
QY	1632	GGAATGATTCAGGTCCTTGGCCAAAGCGAGGCTTGACTGTGAGGGAATGAATG	1691
Db	1498	GGCATGATTCAGGTCCTTGGAAAGTGACGGTGTTCGTGATGTCGAAAAACAACGATG	1557
QY	1692	CCACGATTGGTTATGTTCTAGAGAGAAACGACCAGGCTTAACCATCATAGAAAGCT	1751
Db	1558	CCTCGATTAGTTACGTTTCTCGTGAGAAAGAACCCCGATTGATCACCATAGAAAGCT	1617
QY	1752	GGTCTATGAATGCATGGTCCGAGTCTCTGCTGTACTTAACAAATGCTCCATATTGTTA	1811
Db	1618	GGAGCTATGAATTCCTGTATACGAGTCTCGGGTCTTATCAAAATGCTCCTTACTTCTG	1677
QY	1812	AACTTGATTGTGATCACTACATCAACAACAGCAAGGCTATAAAGAGCAATGTGTTT	1871
Db	1678	AATGTCGATTGTGATCACTACATCAACAATAGCAAGCTTATAGAGAACATGTGTTT	1737
QY	1872	ATGATGAGACCCCTTACTAGAGAAAGAGGTTTGCCTATGTACAGTTCCTCCTCAAGATTGAT	1931
Db	1738	ATGATGATCCTCAGTCAAGAAAGAAATCTGTTATGTTCAAGTTCCTCCTCAAGGTTGAT	1797
QY	1932	GGGATTGATCGCCATGACCGATATGCTAACCGGAATGTGTCTTTTGTATATCAAGT	1991
Db	1798	GGGATTGATAGCAGCAGATCGATACTCAATCGCAATGTGTGTTCTTGTATATCAATATG	1857
QY	1992	AAAGTTTGATGATGATTCAGGGTCCAATTATGTGTAATGATGTGATTTAGAAAG	2051
Db	1858	AAAGTTTGATGGGCTACAGCGCCCTATATACGTCGGTACAGTGTGTGTTTCAGAGG	1917
QY	2052	CAGGCATTATATGTTATGATGCCCCCAAAACAAGAGCCACATCAAGGACTTGCAAC	2111
Db	1918	CAAGCCTTTACGATTGTGATGCCAGAGAGAGAGAGGCCCCACGTAAAGCATGCAAT	1977
QY	2112	TGCTGCCCAAGTGTGCTTTGCTGTGCTGCTTGGCAATAGGAAGCAAAAGAGACT	2171
Db	1978	TGCTGCCCAAAATGTGTCTCTATGTTTGTTCAGAAAGAAATCGTAAAGCAAGACA	2037
QY	2172	ACCAAAACCAAAACAGAGAGAAAAAGTTATTATTTTCAAGAAAGAGAGAACCAATCC	2231
Db	2038	GTGGCTGCGATTAAGAGAGAGAA-----TAGGGAAGCGTCAAGACGATC	2085
QY	2232	CCTGCATATGCTCTTGGTGAATTGACGAAGCTGCTCCAGAGCTGAGATGAAAGGCC	2291
Db	2086	CACGCATTAGAAATATCGAAGAGGGCGCGTCACTAAAGGTTCTAACGTAGAACAGTCA	2145
QY	2292	GGTATTGTAATCAACAAAAATTAGAAAGAAATTTGCCAATCTTCTGTTTGTTACA	2351
Db	2146	ACCGAGGCAATGCAAAATGAGTTGAGAGAAAAATTTGGGCACTCCTGTATTGTGCA	2205
QY	2352	TCCACACTTCTCGAATGTTGGAACCTTGAAGGTGCAAGTCTGCTTCTTTGAAA	2411
Db	2206	TCTGCGCGTATGAGAAATGGTGGGATGGCTAGAAAACGCAACCCGCTGTCTGCTTAAA	2265
QY	2412	GAAAGCTATACATGTCAATTAAGTTGTGTTATGAAGACAAGACAGACTGGGAAAAGAGATT	2471
Db	2266	GAAAGCATCCAAATCATTAAGTTGCCGATATGAAGATTAACCTGAATGGGAAAAGAGATT	2325
QY	2472	GGCTGATCTATGATCAGTTACAGAAAGATATTCTAAGTGTTCAGATGATGATTCAT	2531
Db	2326	GGGTGATCTATGATCTGTATACCGAAAGATATTCTTACGGGTTTAAAGATGATTCAT	2385
QY	2532	GCTTGGCGGTCATTTACTGCAATACCTAAACGGGTTGCAATCAAGGTTCTGACACTCTG	2591
Db	2386	GCTTGGAGATCTGTTTATGTATACCAAAAGTTAGCGGCTTCAAAGATCAGCTCCAATC	2445
QY	2592	AATCTTACAGATCGTCTTACACAGGTGCTTCGGTGGGCTTTGGGCTATATGAGATCTTC	2651
Db	2446	AATCTTTCGATCGTCTTCATCAAGTTCCTTCGATGGGCGCTTGGGTCGGTTGAGATTTTC	2505
QY	2652	TTACGCAATCATGGCCCTCTTTGGTATGCGGTATGCGGTGCGGCTCGAAATTTTGGAAAG	2711

Db	2506	TTGAGTAGGCATTGTCCTATTGGTATGGTTAGAGGTGGGTGAATGGCTTGAGCGG	2565
QY	2712	TTTTCCATCATCACTCCATCGTGTATCCTTGGACATCTATTCGCCCTTGCTTACTGT	2771
Db	2566	TTGTCCTACATTAACTCTGTGGTTTACCCTGGACCTCTACCCGCTCATCGTTTACTGT	2625
QY	2772	ACATTGCCCTGCCATCTGTTTATTGACAGGGAAATTATCACTCCAGAGCTGAATATGTT	2831
Db	2626	TCTCTCCCTGCCATCTGTCCTTCTCACTGGAAAATTTCATCGTTCGCCGATTAGCACTAT	2685
QY	2832	GCCAGCCTGTGCTTCATGTCACTTTTATCTGCAATTTTGTCTACGAGCATCCTAGAAATG	2891
Db	2686	GCGAGTATCCTTCATGAGCGCTCTTCTCGTCGATTGCAATAACGGTATTCTCGAGATG	2745
QY	2892	AGATGAGTGTGTTGGAATTGATGACTGGTGGAGAAATGAGCAGTTCTGGCTCATTTGA	2951
Db	2746	CAATGGGCAAAAGTTGGGATCGATGATTGGTGGAGAAAGAACAGTTTGGGTCAATTGA	2805
QY	2952	GGTGTGTCCTCACACCTCTTGTCTGTGTTCCAGGACCTTCTCAAGTCAATAGCTGTGTT	3011
Db	2806	GGTGTTCCTGCGCATCTGTTGTCTCTTCCAAGTCTCTCAAGTCTTGTGCTGTGTC	2865
QY	3012	GATACAGCTTCACCGTGAACATCAAGGGTGGAGATGATGAGGAGTTCTCAGAGCTATAT	3071
Db	2866	GACACTAATTCAAGTCAACATCAAAAGCAGCTGATGATGAGAGTTCTGTGACCTTTAC	2925
QY	3072	ACATTCAAAATGACTTAACCTTATTGATACCTCTCAACCACTTGCTCTTATTGAACCTTCA	3131
Db	2926	CTCTTCAAAATGACTTCACTTCTCATCTCCCTTCAATGACTCTACTCATATAACGTCA	2985
QY	3132	GGTGTGTCGCTGGCGTTTCAAAATGCGATCAATAACGATATGATCAAGGGCCCTTC	3191
Db	2986	GGAGTCATAGTCGAGTCTCTGATGCCATCAAGCAATGATACGACTCGTGGGACCGCTT	3045
QY	3192	TTTGGGAAGCTATTCTTGCATTTTGGGTGATGTCATCTTTATCCCTTCTCAAGGT	3251
Db	3046	TTCGGAAGACTGTTCTTGCACCTTGGGTCAATTCATCTTACCCGTTCTTAAGGT	3105
QY	3252	TTGGTTGGAAGCAAAACAGACACCAACGATGTCATCGTGTGTCATTTCTGTGCT	3311
Db	3106	TTGCTTGGAAACAAGATAGAATGCCAACCATTAATGTGCTGTGTCATCCTCTGGCC	3165
QY	3312	TCAATCTTCTCGCTCTTGGGTTGGGATGATTCCTTCTTGCGAAGGATGATGCTCG	3371
Db	3166	TCGATTCTTACACTTCTTTGGGTCGGGTTAATCCGTTTGTGGCGA---AAGGCGGTCTT	3222
QY	3372	CTTCTTGAGGAGTGTGGTTTGGAATGC	3398
Db	3223	ATTCTCGAGATCTGTGGTTTGAAGCTGC	3249

```

RESULT 13
US-09-938-842A-559
; Sequence 559, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379

```



; SEQ ID NO 559  
; LENGTH: 3255  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-559

Query Match 36.4%; Score 1431.8; DB 10; Length 3255;  
Best Local Similarity 66.0%; Pred. No. 0;  
Matches 2166; Conservative 0; Mismatches 1052; Indels 63; Gaps 4;

```
QY 144 ATGAGGCGAGCGCGGGCTGGTGCCCGGCTCCACAAACCGCAAGCTCGTCGTCATC 203
Db 1 ATGAATACTGGTGGTCGGCTCATGTGGCTCTCACAACAGAAACGAATTGTTCTCATTT 60
QY 204 CGCCGCGAGCGCGATCCCGGCGGAAGCCGCGGAGAGCAGAACGGGCAAGTGTGCCAG 263
Db 61 AACGCCGATGAGAGTCCAGAAATACGATCACTACAAGAACTGAGTGGCAAAACATGTCAA 120
QY 264 ATTTGCGGCGAGCAGCGTGGCCCTTGCCCGCGGGGAGACCCTTCGTGGCGTGCAACGAG 323
Db 121 ATCTGTGAGATGAATCGAATTACGGTTAGCAGTAGAGCTCTTGTGCTTGCAACGAA 180
QY 324 TGGCCCTTCCCGCTCTGCGCGGACTGCTACGAATACGAGCCGCGGAGGCGACGAGAAC 383
Db 181 TCGGCATTTCCCGGTTTGTAGACCATGCTATGAGTATGAACGTAGAGAGAAATCAAGCT 240
QY 384 TGGCCCCAGTGCAGAACTCGATACAAAGCCCTCAAGGGCTGCCAAGTGTGACCGGTGAC 443
Db 241 TGTCTCAGTGCAAAACCTCGATACAAAAGGATTAAGGTAGTCCACGGGTTGATGAGAT 300
QY 444 GAGGAGGAGAGCGCGTGCATGACCTGGACAAGTTCAACTGGGACGGCCATGACTCG 503
Db 301 GATGAAGAAGAAAGACATTGATGATCTTGAGTATGAGTTTGATCATGGGATGACCCT 360
QY 504 CAGTCTGTGGCCGAGTCCATGCTCTACGGCCATGAGCTACGGCCGTGGAAGTGACCT 563
Db 361 GAACATGCCGCTGAAGCCGCACTCTTCAAGCCTTAACACCGGTCGTGGTGGATTGGAT 420
QY 564 AATGGCGGCGACAAGCTTTCAGCTCAACCCCAATGTTCCACTCTCAACCAACGGGCAA 623
Db 421 TCAGCTCCACCTG-----GCTCTAGATTCTCTTTGACTTATTGTGAT 465
QY 624 ATGTTGATGACATCCACCGGAGCAGCAGCGCTGTCCTTCTTTCATGGGTGTGG 683
Db 466 GAAGATGCTGATATGTATTCTGATCGTCAATGCTCTTATCGTCCCTCTTCAACGGGATAT 525
QY 684 GGAAGAGGATACATCCCTTCCTTATGCGGATCCCACTTACCTGTGCAACCCAGGTCT 743
Db 526 GGGAAATCGCTATCTCTGCAACCGTTTACAGATTCTTCTGCACTCCACAGCGAGATCA 585
QY 744 ATGAGCCCATCCAAGATCTTGCTGATATGGGTATGTTAGTGTGCTTGAAGAAACGG 803
Db 586 ATGTTCTCTCAGAAAGATATTGCGGAATATGTTATGGAAGTGTGCTTGAAGAACCGT 645
QY 804 ATGAGAATTGAGCAGACAGACAAGAGAGATGCAACGAGCGGAATGATGTGTGT 863
Db 646 ATGGAAGTTGGAAGAGACGACAGCAAGGAAAGCTTCAAGTCATTAAAGCATGAAGAGA 705
QY 864 GATGATGT-----GACGATGCTGATTTACCACTAATG 896
Db 706 AACCAATGTCGAGTTCCAATGATGACGACGAAGTATGATCTTGACATGCTTATGATG 765
QY 897 GATGAAGCAGACAAACAATCTGCCAGAAATTCGACTTCATCAAGCCAGATTAAATCCA 956
Db 766 GATGAAGGAAGACAACTCTCTCAAGAAAGCTACCTATTGCTCAAGCAGAAATAATCTT 825
QY 957 TATAGATGATTATCATTTATTCGGCTGTGTGTTGGGGTCTTCTTCACTAACCGAGTG 1016
Db 826 TACAGAGATGTTAATTTCTGTGTGCTCGCGATTCCTGCTTTCTTTCATTATAGAAAT 885
QY 1017 ATGCATCCGTTGAATGATGATTTGCTTGTGGCTCATATCTGTTATCTGAAATCTGG 1076
Db 886 CTCATCCAGTCAATGATGATGATGATTATGTTAAAGTCAAGTATATGCGAGATATGG 945
```

```
QY 1077 TTTGCCATGCTTGGATTCTTGATCAATTCCCAAGGTGTTCCCTATTTGAGAGAGAACT 1136
Db 946 TTTGCAGTGTCTTGGATTCTTGATCAATTCCCAAGGTGTTCCCTATTTGAGAGAGAACT 1005
QY 1137 TACCTAGACCGGCTGTCACTGAGGTTGCAAGGAAGGCCAGCCATCTCAACTGTGCTCA 1196
Db 1006 TACCTGATAGACTCTCTCAGGTACGAAAGGAAGAAAACCGTACAGATTAGCACTT 1065
QY 1197 ATTGATTTCTTTGTACGTAAGGTTGATCCCTTAAAGAACTTCTTGTGTCACAACTAAT 1256
Db 1066 GTTGATGTTTGTGTAGTACAGTGGATCCGTTGAAGAGCCACCCTTGATTACGCAAAAC 1125
QY 1257 ACTGTTCTATCTATCCCTTCGGTGGATTATCCGTGTGATAAGGTTTCTGTATGTTCT 1316
Db 1126 ACAGTTCTTCCATTCTAGCAGTTGATTATCCGTGTGATAAGGTTGCGTGTATGATCA 1185
QY 1317 GATGATGCTGCTGCAATGCTAACGTTTGAAGCATTAATCTGAACATCTGAATTGCAAAAG 1376
Db 1186 GACGATGCTGCACTATGCTTACATTGGAAGCTCTCTGTATACAGCTGAGTTTCTAGA 1245
QY 1377 AATGGGTTCTTCTGCAAAAGGTACAAATATTGAACCTGCGCTCCAGAGTGTACTTC 1436
Db 1246 AATGGGTTCTTCTTGTGAAGATTATATATCGAGCCAGAGCTCCTGAGTGTATTTT 1305
QY 1437 CAACAGAAAGATAGACTACTTGAAGAACAAGGTGGCAGCAAACTTTGTTAGGAGAGAGA 1496
Db 1306 TCTCAGAAAGATGATTACTGAAAGAACAAAGTTCACTCTTGTTCAGGGAAGTGTCT 1365
QY 1497 GCAATGAAGAGAGATGAGGAATTCAAGGTGAGAAATCAATGCTTACTTGGCAAAAGCC 1556
Db 1366 GCTATGAAGAGAGATTATGAGGAGTTAAAGTGAAGATAAATGCACTGTTGCTACTGCA 1425
QY 1557 CAGAAAGTTCTGGAAGAAGATGGAACAATGCAAGATGGAACCCCTGCGCTGGAACAAT 1616
Db 1426 CAGAAAGTGCCTGAGGAGAGTTGGAAGTATGCAAGATGGAACCTTGGCCCTGGAACAAC 1485
QY 1617 GTTGTGATCATCTCGAATGATTTCAAGTCTTCTTGGCCAAAGCGGAGGCTTGACTGT 1676
Db 1486 GTCCGTGACCATCTCGAATGATTTCAAGTGTCTTGGGTCAATGAGATTGCTGATACG 1545
QY 1677 GAGGAAATGAACCTGCCACGATTGTTATGTTCTAGAGAGAAACGACGAGCTAATAAC 1736
Db 1546 GATGTAATGAGTTAACACGCTAGTATGTTCTCGTGAAGAACGCGCTGATTTGAT 1605
QY 1737 CATCATTAAGAAAGCTGTGCTATGAATGATGTCGAGTCTCTGCTGTACTAACAAT 1796
Db 1606 CACCAACAAGAAAGCTGAGCTATGAATCTCTGATCCGAGTCTCTGCTGTCTATCAAAAC 1665
QY 1797 GCTCCATATTGTTAACTTGATGATGATCACTACATCAACAACAGCAAGCTATAAAG 1856
Db 1666 GCTCTTACCTTCTTAATGTGCAATTGTGATCACTACATCAACAACAGCAATTAGA 1725
QY 1857 GAAGCAATGCTTTTATGATGAGACCCCTTACTAGGAAAGAGGTTGCTATGTACAGTTC 1916
Db 1726 GAATCTATGTGTTTCATGATGAGACCCGCAATCGGGAAGAAAGTTGTTATGTTCAGTTT 1785
QY 1917 CCTCAAGATTGATGGGATGATTCGCCATGACCGGATATGCTAAACCGGAATGTGTCTTT 1976
Db 1786 CCGCAGAGATTGATGGGATGATAGACATGATAGATATCAAAACCGTAAACGTTGTCTTC 1845
QY 1977 TTTGATATCAACATGAAAGGTTTGAATGATGATTCAGGGTCAATTATGTGTACTGGA 2036
Db 1846 TTTGATATTAACATGAAAGGCTTTGATGAGATCAAGAGCGGATATATGTGCGGACAGGT 1905
QY 2037 TGTGATTTAGAAGGAGGCAATATATGTTATGATGCCCCCAAAACAAGAGCCACCA 2096
Db 1906 TGTGTGTTAGAAGACAGGCTTTATGTTTATGATGACCAAAAGAAAGAACCA 1965
QY 2097 TCAAGACTTGCAACTGTGGCCCAAGTGTGCTTTGTGCTGTGCTTGGCAATAGG 2156
Db 1966 GGCAAAACCTGTAAGTGTGGCTTAATAGTGTGTTGTGTGTGGGTTGAGAAAGAG 2025
```



QY	544	ACGGCCGTGAGGTGACCTTAATGGCGGCCACCAAGCTTTCCAGCTCAACCCCAATGTTCC	603
Db	639	CCGGTCGTGTGATTGGATTTCAGCTCCACCCTGGCTCT	683
QY	604	CACCTCTCACCAACGGGCAAAATGGTGGATGACATCCACCGAGCAGCACGCGCTGGTGC	663
Db	684	CTCTTTGACTTATTTGATGAAGATGCTGATATGTATTTCTGATCGTCATGCTCTTATCG	743
QY	664	CTTCTTTCATGGGTGGTGGGGGAAAGAGATACATCCCTTCTTATGCGATCCAGCT	723
Db	744	TGCCCTCCTTCAACGGGATATGGGAATCCGCTCTATCCTGCACCGTTTACAGATTTCTCTG	803
QY	724	TACCTGTGCACCCAGGTCTATGGACCCCATCCAGAGATCTTGCTGCATATGGGTATGTA	783
Db	804	CACCTCCACAGGCGGAGATCAATGGTTCTCAGAAAGATATTTCCGGAATATGGTTATGGAA	863
QY	784	GTTGTGCTTGAAGAACGGATGGAGAAATTTGGAAGCAGACAAAGAGAGATGCACCGA	843
Db	864	GTTGTGCTTGAAGAACCTTATGGAACTTTGGAAGACGACAAAGCGAAAGCTTTCAG	923
QY	844	CGGGAATGATGGTGGTGGTGTGATGATGTT-----GACG	876
Db	924	TCATTAAGCATGAAGAGGAAACAATGGTCGAGTTCCAATGATGACGACGAAGTAATG	983
QY	877	ATGCTGATCTACCACTAAATGATGAAGCAAGACAACAACTGTCCAGGAAATTCACCTTC	936
Db	984	ATCCTGACATGCCCTATGATGATGAAGAGACAACTCTCTCAAGAAAGCTACTATTC	1043
QY	937	CATCAAGCCAGATTAATTCATATAGATGATGATTATCATTAATTCGGCTTGTGTTGGGT	996
Db	1044	GTTCAAGCAGAAATAATCTTACAGGATGTTAATTTCTGTGTCGCTCGCGATTCTTGCTC	1103
QY	997	TCTTCTTCCACTACCGAGTGCATCCGGTGAATGATGCATTTGCTTGTGGCTCATAT	1056
Db	1104	TTTTCTTTCATTATAGAAATTCCTCATCCAGTCAATGATGCATATGATTATGGTTAACGT	1163
QY	1057	CTGTATCTGTGAATCTGGTTTGCATGTCTTGGAATCTTGATCAATTCCTCAAGTGT	1116
Db	1164	CAGTTATATGCGAAATATGTTTGACGTCTTGGAATCTTGATCAATTCCTCAAGTGT	1223
QY	1117	TCCCTATTGAGAGAGACTTACCTAGACCGGCTGTCACTGAGTTGCACAAGGAAGGCC	1176
Db	1224	ATCCTATAGAACGTGAACATACCTCGATAGACTCTCTCAGGTACGAGAAGGAAGAA	1283
QY	1177	AGCCATCTCACTTGTCTCAATTGATTTCTTGTCACTACGTTGATCCCTTAAAGAAC	1236
Db	1284	AACCGTCAGGATTAGCACCTGTTGATGTTTTTGTAGTACAGTGGATCCGTTGAAGAGC	1343
QY	1237	CTCCTTTGGTCAACAACAATACTGTTCTATCTATCTTTCGTTGGTGGATTATCCTGTTGATA	1296
Db	1344	CCCCCTGATTACAGCAACAACAGATTCTTCCATTCTAGACAGTTGATTATCCTGTGATA	1403
QY	1297	AGGTTTCTGCTATGTTTCTGATGATGCTGTGCTGCAATGCTAAAGTTGAAGCATTATCTG	1356
Db	1404	AGGTTGCGTTATGTATCAACAACATGGTGCAGCTATGCTTACATTTGAAGCTCTCTCTG	1463
QY	1357	AAACATCTGAATTTGCAAAAGAAATGGGTTCTTCTGCAACGCTACATATTTGAACCTC	1416
Db	1464	ATACAGCTGATTTTGTCAAAAATGGGTTCTTTTGTAGAAAGTTTAATATCGAGCCAC	1523
QY	1417	GCGCTCCAGAGTGTACTTCCAACAGAAGATAGACTACTTGAAGAACAAAGGTGCAGCAA	1476
Db	1524	GAGCTCCTGAGTGTATTTTCTCAGAAGATGATTACCTGAAGAACAAAGTTTATCTCTG	1583
QY	1477	ACTTTGTTAGGAGAGAGACCAATGAAGAGAGATATGAGAAATTCGAAGGTGAGATCA	1536
Db	1584	CTTTTGTCAAGGAAACGTGCTGCTATGAAGAGAGATTATGAAGAGTTTAAAGTGAAGATAA	1643
QY	1537	ATGCTTATAGTGGCCAAAGCCGAGAAAGTTCTGAAAGAGATGGAACAATGCAGATGAA	1596
Db	1644	ATGCACTGGTGTACTGCAAGAAAGTGCCTGAGGAACGTTGGAAGCTATATGCAGATGAA	1703
QY	1597	CCCCCTGGCCTGGAAACAATGTTCTGATCATCTGGAATGATTCAGGTTCTTCTTGCC	1656

Db	1704	CTCCTTGCCCTGGAAACAACGCTCCGTGACCACTCCGTGAATGATTCAAGTGTCTTGGGCTC	1763
QY	1657	AAAGCGGAGGCGCTTGACTGTGAGGGAATGAATGCGCCACGATTGGTTTATGTTCTAGAG	1716
Db	1764	ATAGTGAGATTCTGTGATACGGATGGTAATGATTAACCAACGCTTAGTGTAATGTTTCTCGTG	1823
QY	1717	AGAAACGACCAAGCTATTAACCATCATAGAAAGCTGTGCTATGAATGCATTGGTCCGAG	1776
Db	1824	AGAAAGCGGCTGGATTGTGATCACCACAAGAAAGCTGAGCTATGAATTCCTTGATCCGAG	1883
QY	1777	TCTGTGCTGTACTAACAAATGCTCCATATTTGTTAACTTGATGTGATCACTACATCA	1836
Db	1884	TCTGTGCTGTCTATCAACGCTCCTTACCTTCTTAATGTGATGTGATCACTACATCA	1943
QY	1837	ACAACAGCAAGCTATATAAGAAAGCAATGTTTATGATGAGCCCTTACTAGGAAAGA	1896
Db	1944	ACAACAGCAAGCAATTAAGAAATCTATGTGTTTCATGATGAGCCCGCAATCGGGAAGA	2003
QY	1897	AGGTTGCTATGTACAGTTCCTCCCTCAAGAATTTGATGGATTGATCCGCATGACCGATATG	1956
Db	2004	AAGTTGTATGTTCAGTTTCCGACAGAGATTTGATGGATTGATAGACATGATAGATACT	2063
QY	1957	CTAACCGGAATGTTGCTTTTGTGATATCAACATGAAAGGTTTGATGTGATTCAGGGCTC	2016
Db	2064	CAAAACCGTAACGTTGTCTTTGATATTAACATGAAAGGCTTGATGGGATCAAGGAC	2123
QY	2017	CAATTTATGTGGTACTGGATGTGTATTTAGAAAGGACGACGATTAATGTTATGATGCCC	2076
Db	2124	CGATATATGTGGGACAGGTTGTGTGTTTAAAGAAACAGGCTCTTATGTTGATGACAC	2183
QY	2077	CCAAAACAAAGAACCCACATCAAGGACTTTCACACTGCTGGCCCAAGTGGTCTTTGCT	2136
Db	2184	CAAGAAGAAAGAACCAACAGGCAAAACCTGTAACTGTGGCTAAATGGTGTGTTGT	2243
QY	2137	GTTGCTGCTTGGCAATAGGAAGCAAAAGAACTAACCAACCAAAACAGAGAAGAAA	2196
Db	2244	GTTGGGTTGAGAAAGAAAGATTAACGAAGAACCAACAGATAAGAAACTAACCTAAAG	2303
QY	2197	AGTTATTATTTTCAAGAAAGAGAACCAATCCCTGTCATATGCTCTTGGTGAATTG	2256
Db	2304	AGACTTCA-----AAGCAGATTTCATGCGCTAGAGAATGTGACGAGAG	2345
QY	2257	ACGAAGCTGCTCCAGGAGCTGAGAAATGAAAAGCCGGTATTGTAATCAACAAAATTAG	2316
Db	2346	GTTTATGTCGCCAGTGTCAATGTGAGAGAGATCTGAAGCAACAAATGAAATTGG	2405
QY	2317	AAAAGAAATTTGGCCAACTCTTCTGTTTGTATACACACACTTCTCGAGAAATGTGAA	2376
Db	2406	AGAAGAAGTTGGACAATCTCCGGTTTGTGTCCTGCTGTTCTACAGAACGGTGGAG	2465
QY	2377	CCTTGAAGAGTGCAGTCTGCTCTCTTTTGAAGAAAGCTATACATGTCATTAGTTGTG	2436
Db	2466	TTCCCGCTAACGCAAGCCCCGCATGTTGTTAAGAGAAGCCATTCAAGTTATTAGCTGG	2525
QY	2437	GTTATGAAGACAGACAGACTGGGAAAAGAGATTGGCTGATCTATGATCAGTTACAG	2496
Db	2526	GGTACCAAGATTAACCGAATGGGAAAAGAGATCCGGTGATTTATGATCCGGTACTG	2585
QY	2497	AAGATATTCTAATGCTTTCAAGATGCATTGTCAATGCTTGGCGGCAATTTACTGCATAC	2556
Db	2586	AAGATATCTGACGGGTTTCAAGATGCATTGCCATGAGTAGAGATCTGTGTACTGTATGC	2645
QY	2557	CTAAACGGGTTGCATTCAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAACAG	2616
Db	2646	CTAAGCGTGACGTTTAAAGGATCTGCTCTTAATTAACCTGTACAGATCGTCTTCAACAG	2705
QY	2617	TGCTTCGTTGGGCTCTGGCTCTATGAGATCTTCTTACGAATCATTTGCCCTCTTGGT	2676
Db	2706	TTCTACGTTGGGCTCTGGCTCTGTAGAGATTTCTTGAGCAGACATTTGTCCGATATGTT	2765
QY	2677	ATGGTATGTTGGCGGCTGAAATTTTGAAGATTTTCTTACATCAACTCCATCGTGT	2736



```
Db 2766 ATGCTTATGCTGCTTTAAATGCTTGAGAGATTCTTTACATCAACTCTGCTCT 2825
QY 2737 ATCCTTGACATCTATTCCTCTTGCTTACTGTACATGCTGCCATCTGTTATTGA 2796
Db 2826 ATCCTTGACCTTCACTTCCATGATGCTCTATTGTTCTCTCCCGCGTTGTTACTCA 2885
QY 2797 CAGGAAATTTATCACTCCAGAGTGAATATGTGCCAGCTGTGTTCAATGCTACTTT 2856
Db 2886 CAGGAAATTCATCTCCCTGAGATGAACAACATACGAGGTATTACTCTTCATGCTCATGT 2945
QY 2857 TTAATCTGATTTTGTCTACAGCACTCTAGAAATGAGATGAGTGTGTTGAATTGATG 2916
Db 2946 TCATATCCATAGCAGTAACTGGAATCCTGAAATGCAATGGGAGGTGTCGGAATGATG 3005
QY 2917 ACTGTGAGAGATGAGCAGTCTGGGTCAATGGAGGTGTCTCACACCTCTTGTCTG 2976
Db 3006 ATTGCTGAGAGAAACGAGCAGTTTGGGTAATCGAGGGGCTCTCTCCATCTATTGTCTC 3065
QY 2977 TGTTCAGGGACTTCTCAAGGTCAATAGCTGTGTGATACAAGCTTCACCGTGACATCA 3036
Db 3066 TGTTCAGGTTTGTCTCAAGTTCTAGCCGAGTTAACACGAATTCACAGTCACTTCA 3125
QY 3037 AGGCTGAGATGATGAGAGATTCTCAGACTATATACATTCAATGAGACTACCTTATTGA 3096
Db 3126 AAGCAGACAGATGAGAGCTTCTCTGAGCTTTACATCTTCAAGTGACAACTTTGTTGA 3185
QY 3097 TACCTCTACACCTGTCTTCAATGAACTTCATGTGTGTCTGCTGCGCTTCAATG 3156
Db 3186 TTCTCTCGACAACACTTCTGATCAATTAACATCATTTGAGATTATGTGCGCTTCTGATG 3245
QY 3157 CGATCAATTAACGATATGATGATGAGGCCCCCTCTTTGGGAAGCTATTCTTGCATTT 3216
Db 3246 CCATTAAGCAATGCTATGACTCATGAGGAGCTCTCTTTGGGAGACTTTCTTCTGCTCTT 3305
QY 3217 GGGTATGCTCCATCTTATCCCTTCTCAAAAGTTTGGTTGGAAGGCAAAACAGAGAC 3276
Db 3306 GGGTCAATGTTTCAATTATACCACTTCTCAAGGAATGCTTGGGAAGCAAGCAAAATGC 3365
QY 3277 CAACGATGTGATGCTGTGCTGCTCAATCTGCTGCTCAATCTTCTGCTTGGGTTTC 3336
Db 3366 CTACGATTAATGTGCTGTCTGCTCAATCTTCTAGCTTCACTTGAACACTTGTGGGTCA 3425
QY 3337 GGATGATCTTCTCTGCGAAGATGATGCTCCGCTTCTTGAAGAGTGTGTTGATTT 3396
Db 3426 GAATTAACCGCTTGTGCTA---AAGGGGACCAAGTGTGAGATCTGTGTCTGAATT 3482
QY 3397 G 3397
Db 3483 G 3483
```

## RESULT 15

```
US-10-209-059-45
; Sequence 45, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3725
```

```
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-209-059-45
Query Match 35.0%; Score 1378.4; DB 13; Length 3725;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 2171; Conservative 0; Mismatches 991; Indels 108; Gaps 9;
QY 119 CGCGGAGATGAGGAGGAGAAAGCATGAGGCGGCGGCTGTGCGGCTCCCA 178
Db 154 CGGTGTGTAGAGAGCGCGAGATGCGGCCAAGGGGATGTGCGAGGCTCTCA 213
QY 179 CAACCGCAAGAGCTGCTGTCATCCGCGCGAGCGGCA-----TCCCGGCGGAAGCC 232
Db 214 CAACCGCAAGAGTTCATGATCCGCCAGACGCGGCGAGCGGCTGTCCCGGCTAAGCC 273
QY 233 GCCCGGAGCAGAACGGGCAAGTGTGCCAGATTGGCGGCGAGCGAGCTGCGCTTGCCTCC 292
Db 274 CACGAAGAGTGCGAATGGGCAAGTCTGCCAGATTGTGGCGACACTGTTGGCGTTTCAGC 333
QY 293 CGGCGGGAGCCCCCTGCTGCGGTGCAACGAGTGCCTTCCCGTGTGCGCGGAGTGTCTA 352
Db 334 CACTGTGATGCTTTGTGCTCTGCAATGAGTGTGCTTCCCTGTCTGCGCGCTGTCTA 393
QY 353 CGAATACGAGCGCGGAGGCGACGCAAGACTGCCCCAGTGCAAGACTCGATCAAGCG 412
Db 394 TGAGTACGAGCGCAAGGAGGAACCAATGCTGCGCTCAAGTCAAGACTAGATCAAGAG 453
QY 413 CCTCAAGGCTGCCAAGTGTGACCGGTGACGAGAGAGAGAGCGGCTGATGACTTGA 472
Db 454 ACAGAAAGGTAGCCCTCGAGTTCATGTGATGATGAGAGAGAGATGTTGATGACTTGA 513
QY 473 CAACGAGTCAACTGGGAGCGCCATGACTCGCATGTGTGCGGAGTCCATGCTCTAAGG 532
Db 514 CAATGAATTCAACTATA-----GCAAGGCAATGGGAAGGG 549
QY 533 CCACATGAGTACGCGCGGTGAGGTGACCCCTAATGCGCGCCACAAAGCTTTCAGCTCAA 592
Db 550 CCCAGAGTGCGAGCTTCAAGAGATGACGCTGATGCTTCTCATCTGTGCGCCATGACCC 609
QY 593 CCCCAAT---GTTCACTCTCTCAACCAAGGCAAAAGTGTGATGACATCCACCGAGCA 649
Db 610 ACACCATCGGATTCACGCTTCAAGTGAAGTGAACAAGATATCTGAGAGATCCCTGATGC 669
QY 650 GCAAGCGCTGTGCTCTTTCATAGGTTGGGGGAAGAGATACATCCCTTCTTCA 709
Db 670 ATCCCTGACCGCTATTCTATCCGCACT-----CCAACATCGAGCTA 711
QY 710 TGGGATCCAGCTTACTGTGCAACCCAGGTCTATGGAACCCATCCAAGATCTTGTCTGC 769
Db 712 TGTGATCCAAAGCGTTCCAGTTCTGTGAGATGTGACCCCTCGAAGGACTTGAATTC 771
QY 770 ATATGGGTATGATGTGTCTTGAAGAGAACGATGAGAAATTGGAAGCAGAGACAGA 829
Db 772 CTATGGGCTTAATAGTGTGACTGGAAGAAAGAGTTGAGAGCTGAGGGTTAAACAGGA 831
QY 830 GAGGATGCACGAGCGGGAATGATGTTGTGTGATGATGTG-----873
Db 832 CAAAATATATGTGCAAGTGAATAATATATCCAGAGGCTAGAGAGACATGAGGGGAC 891
QY 874 -----ACGATGCTGATCTAACCACTAATGATGAAGCAACAACAACACTGTCCAGGA 925
Db 892 TGGCTCAATGAGAGAGATATGCAAAATGTTGATGATGACGCTTACCTTGTAGCCGAT 951
QY 926 AATTCACCTTCATCAAGCGAGTTAATCCATATAGATGATTAATCATTTATTCGGCTTGT 985
Db 952 TGTGCCAATTTCTCAACCAAGCTCAACCTTACCGGATAGTAATCATTTCCGCTTAT 1011
QY 986 GGTTTTGGGTTCTTCTTCACTAACCGAGTGAATGATCCGCTGAATGATGATGCTTGT 1045
Db 1012 CATTCTGTCTTCTTCAATATCGTATCAGTCAATGCTGATGCTTATGATTT 1071
QY 1046 GTGCTCATATCTGTATCTGTGAATCTGTTTGCATGTCTTGATTTGATCAATT 1105
```

Db 1072 GTGGTAGTATCTGTTATCTGTGAGGCTGGTTGCTTGCTTGCTGCTTCTAGATCAGTT 1131  
Qy 1106 CCCAAAGTGSTTCCCTATTGAGAGAGACTTACCCTAGACCGGCTGTCACTGAGGTTGCA 1165  
Db 1132 CCCAAAATGATATCCAAATCAACCGTGAGACATATCTGACAGGCTTGCAATTGAGTATGA 1191  
Qy 1166 CAAGGAAGCCAGCCATCTCAACTTGCTCCAAATTGATTTCTTTGTCAGTACGGTTGATCC 1225  
Db 1192 TAGAGAGGAGAGCCATCACAGCTGGCTCCCATTTGATGTCTTTGTCAGTACAGTGATCC 1251  
Qy 1226 CTTAAGGAACCTCCTTTGGTCAACAACAATCTGTTCTATCTATCCTTTCCGTGATTA 1285  
Db 1252 ATTGAAGAACCTCCACTGATCACAGCCAACACTGTTTGTCCATTCTTGCTGTGATTA 1311  
Qy 1286 TCCGTGTGATAAGSTTTCTTGCTATGTTTCTGATGATGCTGTCATGCTAACGTTTGA 1345  
Db 1312 CCTGTGTGACAAAGTGTCACTGATGTTTCTGATGATGCTCAGCTATGCTGACTTTTGA 1371  
Qy 1346 AGCATTAATCTGAAACATCTGAATTTGCAAGAATGGGTTCCCTTTCTGCAACGGTACAA 1405  
Db 1372 GTCTCTCTGAAACTGCGCAATTTGCTAGAAAGTGCGTTCCCTTTTGTAAAGACACAA 1431  
Qy 1406 TATTGAACCTGCGCTCCAGAGTGTACTTCCAAAGAGATAGACTACTTGAAGACAA 1465  
Db 1432 TATTGAACCAAGCTCCAGAAATTTTACTTGTCTCAAAAAATAGATTACCTGAAGACAA 1491  
Qy 1466 GGTGGCAGAACTTTGTTAGAGAGAGAGCAATGAAGAGAGATGAGGAATTCAA 1525  
Db 1492 AATTCAACCTCATTTGTTAAGAAAGAGACCAATGAAGAGAGATGAAGAATTCAA 1551  
Qy 1526 GGTGAGAAATCATGCTTGTAGTGCCAAAGCCAGAAAGTTCCCTGAAGAAGATGACAAT 1585  
Db 1552 AATAAGAAATCATGCTTGTGTTGCCAAAGACAGAAAGTGCTGGAAGGGGTGACCAT 1611  
Qy 1586 GCAAGATGAAACCCCTGGCTGGAACAATGTCGTGATCATCTGGAATGATTCAGGT 1645  
Db 1612 GGCTGATGAACTGCTTGCCCTGGGAATTAACCTTAGGACCATCTGGCATGATTCAGGT 1671  
Qy 1646 CTTCCTTGCCAAAGCGGAGGCTTGACTGTGAGGGAATGAACCTGCGACGATTGGTTTA 1705  
Db 1672 GTTCTTGGGCAAGTGGTGGCTTGACACTGATGAAATGAATTACACGTTGTCTTA 1731  
Qy 1706 TGTTCCTAGAGAAACGACGAGCTATAACCATATAGAAGCTGCTATGATGC 1765  
Db 1732 TGCTCTCGTAAGAAAGAGACGAGCTTTCAGATCACAGAAAGGCTGGTGAATGATGC 1791  
Qy 1766 ATTGSTCCGAGTCTGCTGTACTTAACAATGCTCCATATTTGTTAACTTGATTTGTA 1825  
Db 1792 ACTGATTCGTATATCTGCTGTGACAAATGCTCCTATCTTCTCAATGTGATTTGA 1851  
Qy 1826 TCACTACATCAACAACAGCAAGGCTATAAAGGAAGCAATGTGTTTATGATGAGCCCTTT 1885  
Db 1852 CCATTACTTCAATAGCAAGCAAAAGCTCTTAGAAGCAATGTGCTTCATGATGATCCAGC 1911  
Qy 1886 ACTAGGAAGAGSTTTGCTATGTACAGTTCCTCAAGATTGATGGGATGATGCCA 1945  
Db 1912 TCTAGGAAGGAAAACTGTTATGTACAATTTCACAAAGATTGATGGCATTTGACTTGCA 1971  
Qy 1946 TGACCGATATGCTAACCGGAATGTTGCTTTTGTGATATCAACATGAAGGTTTGATGG 2005  
Db 1972 CGATGATATGCTAATAGGAACATAGTCTTCTTGATATCAACATGAAGGCTTAGATGG 2031  
Qy 2006 TATTCAAGGTCCAATTATGTTGTAAGTGTATTTAGAGGCGGCAATTATATGG 2065  
Db 2032 CATTCAGGTCAGTCTATGTGGAACAGATGCTGTTCAATAGGCGGCTTTGTATGG 2091  
Qy 2066 TTATGATGCCCCCAAAAACAAGAAGCCATCAAGGACTTGCAACTGCTGGCCCAAGTG 2125  
Db 2092 ATATGAT-----CCTGTTTGACTGAAGCTGATCTGGAACCTTAACATT 2134  
Qy 2126 GTGCTTTGCTGTTGCTGCTTTGGCAATAGGAAGCAAAAAGAGACTTAACCAACCAAAAC 2185

Db 2135 GTTGTAAAG---AGCTGCTGTGTAGAAGGAAGAGAAAGAACAAAGATTATATGATAG 2190  
Qy 2186 AGAGAAAGAAAAAGTTATATTTTCAAGAAAGAGAGAACCAATCCCCTGATATGCTCT 2245  
Db 2191 TCAAGCCGTA-----TTATGAGAGAGACAGAAATCTTCAGCTCCCATCTTAACAT 2241  
Qy 2246 TGGTGAATATGACGAAGCTCTCCAGAGCTGAGAAATGAAGGCGCGTATTTGTAATCA 2305  
Db 2242 GGAAGACATCGAGAGGATTTGAAGGTTATGAGGATGAAGAGTCAGTGCTTATGCCA 2301  
Qy 2306 ACAAAAATTGAAAAAGAAATTTGGCCAAATCTTCTGTTTGTTCATCCACACTTCTCGA 2365  
Db 2302 GAGGAATTTGAGAAACGCTTTGGTCAAGTCTCCAATCTTCATTGCATCCACCTTTATGAC 2361  
Qy 2366 GAATGCTGAACCTTGAAGAGTCCAAGTCTGCTTCTCTTTGAAAAGAGCTATACATGT 2425  
Db 2362 TCAAGGTGCATACCACTTCAACAAACCAGCTTCTTACTGAAGGAGCTATTCATGT 2421  
Qy 2426 CATTAAGTTGTTATGAGACAAGACAGACTGGGAAAAAGATTTGCTGAATCTATAG 2485  
Db 2422 TATCAGCTGTGGTACGAGGACAAAACCTGAATGGGAAAAAGATTTGCTGAATCTATAG 2481  
Qy 2486 ATCAGTTACAGAAGATATCTAACTGTTTCAAGATGCAATGTCATGTTGGCGGTCAAT 2545  
Db 2482 TTCAAGTTACAGAGATTTCTGACTGGTTTAAATGCAATGCAAGAGCTGGCAATCAAT 2541  
Qy 2546 TTAATGCATACCTAAACGGGTTGCATTCAAAAGTTCTGCACCTCGAATCTTTGACATCG 2605  
Db 2542 CTACTGCATGCCACCAAGACCTTGTTCAAAGGTTCTGCACCAATCAATCTTTCTGATCG 2601  
Qy 2606 TCTTCAACCAAGGTGCTTGGGCGCTCTTGGGTCTATGAGATCTTCTTACGAATCATTG 2665  
Db 2602 TCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAAGTGAATTTCTGTAGCAGACATTG 2661  
Qy 2666 CCTCTTTGATAGGATAGGTGCGGCTCGAAATTTTGAAGAAATTTTCTTACATCAA 2725  
Db 2662 TCTTATATGATAGCTGCTCAATGGGCGATTGAAGCTTTGAGAGAGCTGCTTACATTA 2721  
Qy 2726 CTGCATGCTATCCTTGGACATCTATTCCTCCCTGCTTACCTGATGCTGCTCAT 2785  
Db 2722 CACCATGTTTATCCAATCAATCTGTCCGCTTATCGCCTATGTGTGCTTCCGTAT 2781  
Qy 2786 CTGTTTATGACAGGAAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGTT 2845  
Db 2782 CTGCTTCTTACCAATAAATTTATCATCTCTGAGATTAGTAATATGCTGGAATGTTCTT 2841  
Qy 2846 CATGTCATTTTATCTGCAATTTTGTCTACGAGCATCTTAGAAATGAGATGAGTGTGT 2905  
Db 2842 CATCTCTCTTTTGGCTCCATTTTGCACACTGTGATATGAGACTCAGATGAGTGTGT 2901  
Qy 2906 TGGAAATGATGACTGTGAGGAATGAGCAGTTCTGGGTCAATGGAGTGTGTCTTACA 2965  
Db 2902 TGGCAATGAAGATTGTGAGAAATGAGCAGTTTGGGTTATTTGGGCACTGTGCCA 2961  
Qy 2966 CTTCTTGTGCTGTGTTCCAGGACTTCTCAAGTCAATAGCTGTGTGATACAAAGCTTAC 3025  
Db 2962 TCTCTTGGCGGTGTTCCAGGCTGTGCTGAAGTGTGGCTGGGATTGATACCAACTTAC 3021  
Qy 3026 CGTGACATCAAAAG---GTGAGATGATGAGAGTTCACAGACTATATACATTCAAAATG 3082  
Db 3022 AGTTAATCTCAAGGATCTGATGAGGATGGCGACTTGTGAGCTATATGTGTCAAGTG 3081  
Qy 3083 GACTACCTTATTTGATACCTCTCAACCACTTGCTTCTTATGAACCTTCATTTGCTGTGCTC 3142  
Db 3082 GACCAATTTGCTCATCCCTCCGACCACTGTTCTTGTCAATTAACCTGTGGAATGTGCGC 3141  
Qy 3143 TGGCGTTTCAAAATGCGATCAATAACGATATGAGTCAATGGGCGCCCTCTTTGGAAAGCT 3202  
Db 3142 AGGAATTTGCTATGCCATTAACAGCGGCTACCAATCTGGGGTCCGCTCTTTGAAAGCT 3201  
Qy 3203 ATTCCTTGCAATTTTGGGTGATGTGCATCTTATCCCTTCTCAAAAGTTGGTTGAAG 3262  
Db 3202 GTTCTTCTGCATCTGGGTGATCTCCATCTCTACCCCTTCTCAAGGCTGTCATGGGCGAG 3261

QY 3263 GCAAAACAGACACCAACGATGTCATGCTGGTCATCTGCTGCTTCAATCTTCTC 3322  
DB 3262 GCAGAACCGCACGCGCAACATCGTCATCGTTGGTCCATCCTTGCGTCTATCTTCTC 3321  
QY 3323 GCTCCTTTGGGTTCCGATGATCCTTCTC 3352  
DB 3322 CTTGCTGTGGGTGAAGATGATCCTTTCAT 3351

Search completed: December 16, 2003, 04:38:54  
Job time : 1289 secs





Db 459 GTAGAGAAAGAAATCAAGCTTGTCTCAGTGCAAAACTCGATACAAAAGATTAAAGTA 518  
Qy 424 GCCAACGTGTGACCGGTGACGAGAGAGAGACGGCGTCGATGACCTGACAAAGATTCA 483  
Db 519 GTCCACGGGTTGATGAGATGATGAAGAAGAGAATGATGATCTTGAGTATGAGT 578  
Qy 484 ACTGGACGGCCATGACTCCAGTCTGTGGCCGAGTCCATGCTCTACGGCCACATGAGCT 543  
Db 579 TTGATCATGGGATGAGACCTGAAACATGCCCGTGAAGCCGCACCTCTTCAAGCCTTAACA 638  
Qy 544 ACGGCCGTGAGGTGACCTTAATGGCGGCCACAAAGCTTTCAGCTCAACCCCAATGTTT 603  
Db 639 CCGGTGCTGATGATGATGATCAAGCTCCACCTGGCTCT-----CAGATTTC 683  
Qy 604 CACTCCTCACCACGGCAATGTGTGATGACATCCACCGAGCAGACGGCGTGTGC 663  
Db 684 CTCTTTGACTTATGTGATGAAGATGCTGATATGTATTTCTGATCGTCATGCTCTTATCG 743  
Qy 664 CTTCCTTTCATGGGTGGGGGAAAAGAGATACATCCCTTCTTATGCGGATCCCAAGCT 723  
Db 744 TGCCCTCCTCAACGGGATATGGGAATCGCGTCTATCCTGCAACCGTTACAGATTCTTCTG 803  
Qy 724 TACCTGTGCAACCCAGTCTATGGAACCCCATCCAAAGATCTTGCTGATATGGGTATGTA 783  
Db 804 CACCTCCACAGCGGAGATCAATGGTTCCTCAGAAAGATATTCGGGAATATGGTTATGGA 863  
Qy 784 GTGTGCTTGAAGGAAACGATGGAATTTGAAGCAGACAAAGAGAGATGCAACAG 843  
Db 864 GTGTGCTTGAAGGACCGGTATGGAAGTTTGAAGAGACGACAAAGCGAAAGCTTCAAG 923  
Qy 844 CGGGGAATGATGGTGGTGTGATGATGTT-----GACG 876  
Db 924 TCATTAAAGCATGAAGAGAGAAACAATGTCGAGGTTCCAATGATGACGACGAAGTAAGTG 983  
Qy 877 ATGCTGATCTACCACTAATGATGAAGCAAGACAAACAAGTGTCCAGAAAATTCACCTTC 936  
Db 984 ATCCTGACATGCCCTATGATGATGAAGAGACAACTCTCTCAAGAAAGCTACCTATTC 1043  
Qy 937 CATCAAGCCAGATTAATCCATATAGCATGATTATCATTTATCCGCTGTGTTGGGT 996  
Db 1044 GTTCAAGCAGAAATAATCCTTACAGCATGTTAATCTGTGTCCGCTCGCGATTCTTGCTC 1103  
Qy 997 TCTTCTTCCACTACCGAGTATGCATCCGGTGAATGATGATTTGCTTGTGCTCATAT 1056  
Db 1104 TTTTCTTTCATTATAGAAATCTCCATCCAGTCAATGATGATGATTATGTTAAAGT 1163  
Qy 1057 CTGTTATCTGGAATCTGGTTTGCCATGTCTTGATCTTGATCAATTCCTAAAGTGT 1116  
Db 1164 CAGTTATATGCGAAATATGTTGACGTCTTGATCTTGATCAATTCCTAAATGTT 1223  
Qy 1117 TCCCTATTGAGAGAGAGACTTACCTAGACCGGCTGCACTGAGGTTGACAAGAAAGCC 1176  
Db 1224 ATCCTATAGAACGTGAACATACCTCGATAGACTCTCTCAAGTACGAGAAGAAAGAA 1283  
Qy 1177 AGCCATCTCAACTTGTCCCAATGATTTCTTGTGCAATACGTTGATCCCTTAAAGAAC 1236  
Db 1284 AACCGTCAGGATTAGCACCTGTGATGTTTGTAGTACAGTGATCCGTTGAAGAGC 1343  
Qy 1237 CTCTTTGTCACAACAATACTGTTCTATCTATCTTTCGGTGATATCTGTTGATA 1296  
Db 1344 CCCCCCTGATTACAGCAAAACAGTCTTTCATTCTAGCAGTTGATTATCCGTGATA 1403  
Qy 1297 AGGTTTCTGCTATGTTTCTGATGATGCTGCTGCAATGCTAACGTTGAAGCATTATCTG 1356  
Db 1404 AGGTTGCGTGTATGATCAAAACAATGCTGACGCTATGCTTAATTGAAGCTCTCTCTG 1463  
Qy 1357 AAACATCTGAATTTGCAAAAGAAATGGGTTCTTCTGCAAAAGGTAACAATATGAACCTC 1416  
Db 1464 ATACAGCTGATTTTGTCTACAAATAGGGTCTTTTGTAGAGAGTTAATATCGAGCAC 1523  
Qy 1417 GCGCTCCAGAGTGTACTTCCAAAGAGATAGACTACTTGAAGAAGCAAGGTGCGACAA 1476  
Db 1524 GAGCTCCTGAGTGTATTTTCTCAGAGATGATTACCTGAAGAACAAGTTCAATCCTG 1583

Qy 1477 ACTTTGTAGGAGAGAGAGCAATGAAGAGAGATATGAGAAATTCAAAGGTGAGATCA 1536  
Db 1584 CTTTGTACGGGAACGTCGTCTATGAAGAGAGATTATGAAGATTTAAAGTGAAGATAA 1643  
Qy 1537 ATGCTTAGTTGCCAAAGCCCAAGAGTTCTGAAAGAGATGACAATGCAAGATGGA 1596  
Db 1644 ATGACCTGGTTGCTACTGCAAGAAAGTCCCTGAGGAACGTTGACTATGCAAGATGAA 1703  
Qy 1597 CCCCCGTGCGAAACAATGTTGATGATCATCCTGGAATGATCAAGTCTTCTTGCC 1656  
Db 1704 CTCCTTGCGCTGGAACAACGTCGTCACCATCCTGGAATGATTCAAGTGTCTTGCGTC 1763  
Qy 1657 AAAGCGAGGCTTGACTGTGAGGAAATGAAGTCCACGATGTTATGTTCTAGAG 1716  
Db 1764 ATAGTGAGTTGCTGATACGGAATGTAATGAGTTACACGTCCTAGTATGTTCTCGTG 1823  
Qy 1717 AGAAAGCAGCAGCTATTAACCATCATTAAGAAAGCTGCTGCTATGAATGCATTGTCGAG 1776  
Db 1824 AGAAGCGGCTGGATTTGATCACACAGAAAGCTGAGCTATGAATTCCTTGATCCGAG 1883  
Qy 1777 TCTCTGCTGACTAACAAATGCTCCATTTGTTAACTTGATGATGATCACTACATCA 1836  
Db 1884 TCTCTGCTGTTCTATCAAAACGCTCCTTACCTTCTTAATGTCATGATCACTACATCA 1943  
Qy 1837 ACAACAGCAAGCTATTAAGGAAGCAATGTTTATGATGACCTTTACTAGAAAGA 1896  
Db 1944 ACAACAGCAAGCAATTAAGAAATCTATGTGTTTCATGATGAGACCCGCAATCGGAAAGA 2003  
Qy 1897 AGTTTGTATGTACAGTTCCTCCCAAGATTTGATGGATGATGCGCATGACCGATATG 1956  
Db 2004 AAGTTGTATGTTCAGTTTCCCGCAGAGATTTGATGGATTGATGACATGATAGATACT 2063  
Qy 1957 CTAAACCGAATGTTGCTTTTGTGATTAACAATGAAGGTTTGAATGTTTCAAGGTC 2016  
Db 2064 CAAACCGTAACGTTGTGTTCTTGTATTAACAATGAAGGCTTGTGATGCAAGGAC 2123  
Qy 2017 CAATTATGTTGCTACTGATGTGATTTAGAAAGCAGGCAATTAATGTTATGATGCC 2076  
Db 2124 CGATATATGTGGGACAGGTGTGTGTTAGAAAAACAGGCTCTTAATGCTTTGATGCAC 2183  
Qy 2077 CCAAAACAAAGAACCAACCATCAAGGACTTGCAACTGCTGCGCCAAAGTGTCTTTGCT 2136  
Db 2184 CAAAGAAAGAAACCAACGCAAAACCTGTAACGTGTGCGCTAAATGTTGTTGT 2243  
Qy 2137 GTTGCTGCTTGGCAATAGGAAGCAAAAGAAAGCTAACCAAAACCAAAACAGAGAAGAAA 2196  
Db 2244 GTTGTGGTTGAGAAAGAGAGTAAACGAAGCCACAGATAAGAAACTTAACACTAAAG 2303  
Qy 2197 AGTTATTTATTTCAAGAAAGAGAACCAATCCCTGCATATGCTCTTGATGAATTG 2256  
Db 2304 AGACTTCA-----AACGATTTCAATGCGCTAGAGATGTCAGCAAG 2345  
Qy 2257 ACGAAGCTGCTCAGGAGCTGGAATGAAGAGCCGGTATTGTAATCAACAAAATTAG 2316  
Db 2346 GTGTTATGTCGACAGTGTCAATGTGGAAGAGATCTGAAGCAACAATTGAATTGG 2405  
Qy 2317 AAAAGAAATTTGGCCAATCTTCTGTTTGTTAACATCCACACTTCTCGAAGATGTGAA 2376  
Db 2406 AGAAGAAATTTGCAATCTCCGTTTTCGTTGCTCTGCTGTTCTACAGAACGCTGAG 2465  
Qy 2377 CCTGAAGAGTGAAGTCTGCTTCTCTTTGAAGAGCTATACATGTCATTAGTTGTG 2436  
Db 2466 TTCCCGGTACGCAAGCCCGCATGTTGTTAAGAGAGCCATTCAAGTTATTAAGTCGCG 2525  
Qy 2437 GTTATGAAGACAGACAGACTGGGAAAAGAGATTGGCTGATCTATGATCAAGTTACAG 2496  
Db 2526 GGTACCAAGATTAACCGAATGGGAAAAGAGATCGGGTGATTTATGATCGGTGACTG 2585  
Qy 2497 AAGATATTTAATGTTTCAAGATGATGATGTTGCGGTTCAATTTACTGCATAC 2556  
Db 2586 AAGATATCTGACGGGTTTCAAGATGATGATGATGATGAGATCTGTGTAATGCTGATGC 2645

QY	2557	CTAAACGGGTTGCATTCAAAGGTTCTGCACCTCTGAACTTTTCAGATCGTCTTCACCAGG	2616
Db	2646	CTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTATTAACTTGTCAAGATCGTCTTCAACAG	2705
QY	2617	TGCTTCGGTGGGCTCTTGGGCTATTTAGATCTTCTTGAGCAATCATTTGCCCTCTTTGGT	2676
Db	2706	TTCTACGTTGGGCTCTTGGCTCTGTAGAGATTTTCTTGAGCAGACATTTGTCCGATATGGT	2765
QY	2677	ATGGGTATGTTGGCGGCTGTGAATTTTGGAAAGATTTTCTTACATCAACTCCATCGTGT	2736
Db	2766	ATGCTTATGTTGGTGGTTTAAATGTTGGAGAGATTTCTTACATCAACTCTGTCTGTCT	2825
QY	2737	ATCCTTGACATCTATTCCCTCTTGCTTACTGTATCATTTGCCATCTGTTTATTTGA	2796
Db	2826	ATCCTTGACCTTCACTTCCATTTGATCGTCTATTGTTCTTCTCCCGGTTTGTACTCA	2885
QY	2797	CAGGAAATTTATCACTCCAGAGCTGAATTAATGTTGCCAGCCTGTGTTTCAATGTCATTT	2856
Db	2886	CAGGAAATTCATCGTCCCTGAGATAAGCACTACGAGGTATACTCTTCATGCTCATGT	2945
QY	2857	TTATCTGCATTTTGTCTACGACATCTCTAGAAATGAGATGGAAGTGTGTTGAATTTATG	2916
Db	2946	TCATATCCATAGCAGTAACGTGAATCTCTGAAATGCAATGGGAGGTGTGGAATTCATG	3005
QY	2917	ACTGCTGAGGAATGAGCAGTTCTGGGTCATTTGAGAGGTGTCTCACACCTCTTTGCTG	2976
Db	3006	ATTGCTGAGAAACGAGCAGTTTGGGTAATCGAGGGGCTCTCGCATCTATTGCTC	3065
QY	2977	TGTTCCAGGACTTCTCAAGTTCATAGCTGTGTTGATTAACAAGCTTCAACGTCATCA	3036
Db	3066	TGTTTCAAGTGTGCTCAAGTTCAGCCGAGTTAACAGAAATTTCAAGTCACTTCA	3125
QY	3037	AGGCTGAGATGATGAGAGTTCTCAGAGCTATATACATTTCAATGAGCTACCTTATTTGA	3096
Db	3126	AAGCAGCAGACGATGAGCTTTCTCTGAGCTTTACATCTTCAAGTGAGCAACTTTGTTGA	3185
QY	3097	TACCTCCTACCACTTGTCTTCTATTGAATTCATTTGTGTGTGCTGCGCTTCAATG	3156
Db	3186	TTCTCCGACACACTTCTGATCATTTACATCATTTGAGTTATTGTGCGCTTCTGATG	3245
QY	3157	CGATCAATAACGATATGAGTCATGCGGCCCTCTTTGGGAAGCTATTCTTGCATTTT	3216
Db	3246	CCATTAGCAATGCGTATGACTCATGGGACCTCTTTGGGAGACTTTCTTCTGCTCTT	3305
QY	3217	GGGTGATTTGCCATCTTATCCCTTCTCMAAGTTTGGTTGGAAGCAAAACAGACAC	3276
Db	3306	GGGTCAATTGTTCAITTAACCATTCTCTCAAGGGAATGCTTGGGAAGCAAAATGC	3365
QY	3277	CAAGGATTTGATCGTCTGATTCATTTCTGCTGCTTCAATCTTCTGCTCTTTGGGTTT	3336
Db	3366	CTACGATTATTGTGTGTGTCTATTCTTCTAGCTTCGATCTTGACACTCTTTGGGTCA	3425
QY	3337	GGATTGATCCTTTCTCTGGAAGGATGATGTCGCTTCTTGAGAGGTGTGTTGGATT	3396
Db	3426	GAATTAACCCGTTGTGCTA--AAGGGGAGCCAGTGTGAGAGATCTGTGTCTGAATT	3482
QY	3397	G 3397	
Db	3483	G 3483	

RESULT 2  
US-09-221-013A-11  
; Sequence 11, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A

```

; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3673
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (71)..(3313)
US-09-221-013A-11

Query Match      33.8%; Score 1330.8; DB 4; Length 3673;
Best Local Similarity 64.9%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 1067; Indels 75; Gaps 7

QY 127 GTGAGGGGAGGAGGCGATGAGGCGGCGCGGCTGTGCGCGGCTCCCAACCGCA 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 GTGTGTGGTGGCTGCGATGAGGCGCAAGTCCGCGCTTGTTGTGATCCTACCGAGAA 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 187 ACGAGCTCGTCGTCATCCGCCGCGACGGCGATCCCGGCGGAGGCCCGCGGAGCAGA 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 ACGAGCTCGTCGATCCGACATGAATCTGATGCGCGGACCAACCTTTGAAGAATATGA 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 247 ACGGCGAGGTGTGCCAGATTTCGCGCGACGACGTCGCGCTTGCCCCCGGCGGACCCCT 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 ATGCGCCAGATATGTCAAGATCTGTGTGATGATGATGTTGACTCGCTGAAACTGAGATGCT 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 307 TCGTGGCGTGCAACGAGTCCGCTTCCCGTCTGCGCGGACTGCTAAGAAATACGAGCGCC 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 TTGTCCGCTGTAATGAATGTGCTTCCCTGTGTGTGCGCGCTTGCTATGAGTACGAGAGA 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 367 GGGAGGGGACGAGAACTGCCCCAGTGCAGAACTCGATACAAAGCGCCTCAAGGGCTGCC 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 AAGATGGAAGTCAAGTGTGCGCTCAATGCAAGACTAGATTCAAGCAGACAGGGGAGTGC 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 427 AACGTGTGACCGGTGACGAGAGGAGGACGGCGTGCATGACTTGACCAACGAGTTCAACT 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 CTGCTGTGAAGAGATGAAGATGAGGATGATGTTGATGATATCGAAGATGAGTCA--- 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 487 GGGACGGCCATGACTCGCAGTCTGTGCGCGAGTCCATGCTCAACGCCACATGAGTACG 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 -----ATTACGCCCAGGGAGCTTAACAAGGCGAGACACCAACGCCA---TGGCGAAG 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 547 GCCGTGAGGTGACCCCTAATGGCGGCCACAGCTTTCAGCTCAACCCCAATGTTCCAC 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 AGTTTCTCTCTCTCTAGACATGAATCTCAACCAATTCCTCTTCTCACCCATGGCCATA 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 607 TCCTCACCAACGGGCAAAATGCTGATGACATCCCAACCGGAGCAGACGCGCTGTGCTT 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 CGGTTTCTGAGAGATTCCGACGCCCTGATACACAATCTGTGCGAACTACATCAGGTCCTT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 667 CTTTCATGGGTGGTGGGGAAGGATACATCCCTTCTTAATGCGGATCCAGCTTAC 726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 -----TGGTCTTCTGACAGGAATGCTATTTCATCTCCATATATGATCCACGGCAAC 632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 727 CTGTGCAACCCAGGTCTAATGACCCCATCCAGAGATCTTGCTGATATGGGTATGTAAGT 786
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 633 CTGTCCCTGTAAAGATCGTGACCCCGTCAAAAGACTTGAACCTTATGGGCTTGTAATG 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 787 TTGCTTGGAAGAACGATGAGCAATTGGAAGCAGAGACAGAGAGATG----- 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 693 TTGACTGGAAGAAAGAGTTGAAGGCTGGAAAGCTGAAGCAGAGAGAAAAATATGTTACAGA 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 837 -----CACAGACGGGGATGATGTTGCTGTGATGATGATGATGATGATGATGATG 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 753 TGACTGTTAATATACATGAGAGGAGAGAGAGAGAAATGAGGACTGTTTCCAATGGCG 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 880 CTGATCTACCACTAATGATGAAGCAGAGCAACAACTGTCTCAGGAAAAATTCACCTTCAT 939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



Db 813 AAGAACTCCAAATGCTGATGATACACGCTCTTCTATGAGTCGTGTGCTGCTATCCCAT 872  
Qy 940 CAAGCCAGATTAAATCCATATAGATGATATATATATTCGGCTGTGTTTGGGGTCT 999  
Db 873 CTCTCGCCCTAACCCCTTATCGGGTGTGATTATCTCCGGCTTATCATCTTGTGTTCT 932  
Qy 1000 TCTTCCACTACCGAGTGATGCATCCGGTGAATGATGCATTTGCTTGTGCTCATATCTG 1059  
Db 933 TCTGCAATATCGTAACTCACTACCCCTGTGAAAAATGCATATCCTTGTGTGACCTCG 992  
Qy 1060 TTATCTGTGAATCTGCTTGGTCCATGTCTTGATTTCTTGATCAATTCGCAAGTGTTC 1119  
Db 993 TTATCTGTGAGATCTGTTGATTTGCAATTTCTTGGCTTCTTGATCAGTTCCCAATGTACC 1052  
Qy 1120 CTATTGAGAGAGACTTAACCTAGACCGGCTGTCACTGAGGTTGACAAAGAGCCAGC 1179  
Db 1053 CCATTAACAGGAGACTTATCTTGACCGTCTCGCTATTAAGATATGATCGAGACGGTGAA 1112  
Qy 1180 CATCTCAACTGTCTCCAATTGATTTCTTGTGACGAGGTTGATCCCTTAAAGAACCTC 1239  
Db 1113 CATCACAGCTCGTTCCTGTGATGTGTGTTGTTAGTACAGTGACCATTGAAAGACCTC 1172  
Qy 1240 CTTTGTCAACAATACTGTCTATCTATCTTCCGTGCGTGATTAATCTGTGATAAG 1299  
Db 1173 CCCTGTACAGCAACACAGTTCTCTGATTTCTTCTGTGACTACCGGTAGATAAG 1232  
Qy 1300 TTTCTGTATGTTCTGATGATGTGCTGCAATGCTAACGTTGAAGCATTAATCTGAAA 1359  
Db 1233 TAGCCTGTATGTTTACAGATGATGTTTACAGCTATGCTTACCTTGAATCCCTTCTGAAA 1292  
Qy 1360 CATCTGAATTTGCAAGAAATGGGTTCCCTTCTGCAAAACGGTACAATATTTGAACCTCGCG 1419  
Db 1293 CCGTGAGTTTGCAAGAAATGGGTACCATTTTGCAAGAAATTCACACATTTGAACCTAGGG 1352  
Qy 1420 CTCACAGTGTACTTCCAAAGAGATAGACTTGAAGAACAAGGTGCGACGAAACT 1479  
Db 1353 CCCCCTGAATTTCTATTTGCCCCAGAAAGATAGTACTTGAAGACAAGATCCAACCGTCTT 1412  
Qy 1480 TTGTTAGGAGAGAGACAATGAGAGAGATGAGAAATTCAGGTGAGAAATCAATG 1539  
Db 1413 TTGTTAAAGAGCGAGAGCTATGAAGAGAGATGAAGATTAAAGTGAGATAAATG 1472  
Qy 1540 CCTTAGTTGCCAAAGCCCAAGAAAGTTCTCTGAAGAAGATGACAATGCAATGGAACCC 1599  
Db 1473 CTCTGTGTCCAAAGCAGAGAAATCCCTGAGAGAGCTGAGCAATGAGATGTAATC 1532  
Qy 1600 CCTGGCCTGAAACAATGTTCGTGATCATCTGGAATGATTCAGGTCTTCTTGCCCAA 1659  
Db 1533 CCTGGCCTGTAAACAACAATAAGATCATCTGGAATGATACAGGTGTTCTTAAGCCATA 1592  
Qy 1660 GCGGAGGCTTGACTGTGAGGGGAAATGAACCTGCCAGATGGTTATGTTCTAGAGAGA 1719  
Db 1593 GTGGGGTCTGGATACCGATGGAAATGAGCTGCTAGACTCATCTATGTTCTCGTGAAA 1652  
Qy 1720 AACGACCAAGCTATTAACCATATAGAAGAGCTGTGCTATGAATGACATGCTGCGAGTCT 1779  
Db 1653 AGCGCCTGATTTCAACCACAAAAAAGGCTGAGCTATGAATGACATGATCCGTGAT 1712  
Qy 1780 CTGCTGACTAACAAATGCTCCATATTTGTTAACTTGATTTGATGATCACTACATCAACA 1839  
Db 1713 CTGTGTCTTACCAATGAGCATATCTTTGAACGTGATGTGATCATTAATTAAATA 1772  
Qy 1840 ACAGCAAGCTATAAGAGAAAGCAATGTGTTTATGATGAGACCCCTTACTAGAAAGAGG 1899  
Db 1773 ACAGTAAGGCTATTAAGAAAGCTATGTGTTTCAATGGAACCCGGCTATTTGAAAGAGT 1832  
Qy 1900 TTGCTATGTACAGTTCCTCAAGAGATTGATGAGATGATGCCATGACGATATGCTA 1959  
Db 1833 GCTGCTATGTCCAGTTCCTCAACGTTTGAACGATGATTTGACAGATGATATGCCA 1892  
Qy 1960 ACCGAATGTGTCTTTTGTGATATCAACATGAAGGTTTGATGTATTCAGGGTCCAA 2019

Db 1893 ACAGAAATATAGTCTTTTTCGATATTAACATGAAGGGGTGGATGCTATCCAGGTCAG 1952  
Qy 2020 TTTATGTTGGTACTGATGTGTAATTTAGAAGGACGACATTAATGCTTATGATGCCCCA 2079  
Db 1953 TATATGCGGTACTGCTGTGTTTAAATAGGACGCTCTATATGGGTATGATCCTGTTT 2012  
Qy 2080 AAACAAGAACCCACCATCAAGACTTGCAACTGCTGGCCCAAGTGCTTTTGTCTGTT 2139  
Db 2013 TGACGGAAGAGATTAGAACCAAAATATTA-----TTGTCAAGA 2051  
Qy 2140 GCTGCTTTGGCAATAGGAAGCAAAAGAACTATCCAAACCAAAACAGAGAAAGAAAGT 2199  
Db 2052 GCTGTTGCGGTCAGAGAAAGGTAAGAAAGTAGCAA--GAAGTAACTACGAAAAAG 2108  
Qy 2200 TATTAATTTTCAAGAAAGAGAACCAATCCCTGCAATATGCTTGTGTAATTGACG 2259  
Db 2109 GAGAGGCATCAACAGAGTGACTCCAATGCTCCACTTTTCAATATGAGAGACATCGATG 2168  
Qy 2260 AAGCTGCTCAGGAGCTGAGATGAAGAAAGCCGCTATTTGTAATCAACAAAATTAGAAA 2319  
Db 2169 AGGCTTTGAAGTTATGATGATGAGAGGCTATTCTAATGTCCCAAGAGAGTGTAGAG 2228  
Qy 2320 AGAAATTTGGCCAAATCTTCTGTTTGTTCATCCACACTTCTCGAAGATGTGGAACCT 2379  
Db 2229 AGCGTTTGTGTCAGTCGCCGCTATTTATTTGCGCAACCTTCATGGAACAAGCGGCAATTC 2288  
Qy 2380 TGAAGAGTCAAGTCCCTGCTTCTTTTGAAGAAAGCTATACATGTCATTAATGTGTGTT 2439  
Db 2289 CACCAACAACCAATCCCGCTACTCTTCTGAAGAGAGGCTATTCATGTATTAAGCTGTGTT 2348  
Qy 2440 ATGAAGCAAGACAGACTGGGAAAAAGAGATTGGCTGATCTATGATCAGTTACAGAA 2499  
Db 2349 ACGAAGCAAGACTGAATGGGCAAGAGATTGTTGATCTATGTTCCGTGACGGAAG 2408  
Qy 2500 ATATTTAATGTTTCAAGATGATGTGCATGTTGGCGGTCAATTTACTGCATACCTA 2559  
Db 2409 ATATTTCTACTGGTTCAAGATGATGCCCGGGTGGATTCGATCTACTGCATCCTC 2468  
Qy 2560 AACGGTTGCATTCAAAGGTTCTGCACTCTGAATCTTTCAGATCGTCTTCACCAGGTGC 2619  
Db 2469 CACGCCCTGCGTTCAAGGATCTGCAACCAATCAATCTTTCGATCGTTTGAACCAAGTTC 2528  
Qy 2620 TTGGTGGGCTCTTGGGTCTATGAGATCTTCTTACGCAATCATTGCCCTCTTGGTATG 2679  
Db 2529 TTGATGGGCTTTGGGATCTATCGAGATTTCTTCTTAGCAGCAATGTCCTATCTGGTATG 2588  
Qy 2680 GGTATGTGGCGGTCTGAATTTTGAAGAAATTTTCTACATCAACTCCATCGTATATC 2739  
Db 2589 GTTACCATGAAAGGTTGAGACTTTGAGAGAGATCGCTTATATCAACACCATCGTATATC 2648  
Qy 2740 CTTGACATCTATTTCCCTCTTGGCTTACTGTACATCATGCTGCAATCTGTTATGACAG 2799  
Db 2649 CTATTACATCCATCCCTCTTATTTGGTATTTGATTTCTTCCGCTTTTGTCTCATCACCG 2708  
Qy 2800 GGAATTTATCACTCCAGAGCTGAATATGTTGCCAGCCTGTGTTCAATGCACTTTTAA 2859  
Db 2709 ACAGATTCAATCAACCCAGATTAAGCAACTACGCGAGTATTTGTTCAATTTCTACTCTTCA 2768  
Qy 2860 TCTGCATTTTGTCTACGAGCATCTTGAATGAGATGAGTGTGTTGAATGATGACT 2919  
Db 2769 TCTCAATGCTGTGACTGGAATCTGAGCTGAGATGAGCGGTGTGAGCAATGAGGAT 2828  
Qy 2920 GGTGAGGAATGAGCAGTTCTGGGTCAATGGAGGTGTCTTCAACACTCTTGTGTGTG 2979  
Db 2829 GGTGAGGAACGAGCAGTTCTGGGTCAATGGTGACATCCGCCCATCTTTTGTGTGTCT 2888  
Qy 2980 TCCAGGACTTCTCAAGGTCAATAGCTGTGTTGATACAAGCTTCAACCGTGACATCAAGG 3039  
Db 2889 TCCAAGTCTACTTAAGGTTCTGTGCTGATCGACACCAACTTCAACCGTTACATCTAAAG 2948  
Qy 3040 ---GTGAGATGATGAGAGTTCAGAGCTATATACATTTCAATGGACTACCTTATGA 3096  
Db 2949 CCACAGACGAGATGGGATTTTGCAGAACTCTACATCTTCAATGGAAGAGCTCTTCTCA 3008

QY 3097 TACCTCCTACCACCTTGCTTCTTAATTGAACCTTCATTGGTGTGGTCGGCTGGCTTCAATG 3156  
 Db 3009 TTCCACCAACCAACCGTCTTACTGTGTGAACCTCATAGGCAATGTGGCTGTGTCTTTATG 3068  
 QY 3157 CGATCAATTAACGGATATGAGTCATGCGGGCCCCCTCTTGGGAAGCTATCTTGCATTTT 3216  
 Db 3069 CTGTAAACAGTGGCTACCAAGTCGTGGGTCGCTTTTCGGGAAGCTCTTCTGCTTAT 3128  
 QY 3217 GGGTGAATTGCCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGGCAAAAAGAGAC 3276  
 Db 3129 GGGTATATTGCCCATCTCTACCTTTCTTGAAGAAGTCGTGTGGGAAGACAAAACGAC 3188  
 QY 3277 CAACGATGTGCATCGTCTGTGCCATTCTCTGCTGGCTTCAATCTTCTGCTCCTTGGGTT 3336  
 Db 3189 CAACCATCGTCATTTGTCTGTCTGTCTTCTCTGCGCTCCATCTTCTCGTTGCTTTGGGTC 3248  
 QY 3337 GGATTGATCCTTCTCTTG 3354  
 Db 3249 GGATCAATCCCTTTGGG 3266

### RESULT 3

```

US-09-221-013A-5
; Sequence 5, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3603
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3243)
US-09-221-013A-5

```

Query Match	33.7%	Score 1326.6;	DB 4;	Length 3603;
Best Local Similarity	65.0%;	Pred. No. 0;		
Matches 2107; Conservative	0;	Mismatches 1059;	Indels 75;	Gaps 7;

QY 144 ATGAGCGCAGCGCCGGGCTGTGTGCGCGCTCCACAACCGCAACGAGCTGTGTCATC 203  
 Db 1 ATGAGGCCAGTGC CGGCTTGTGTGTGATCTTACCGAGAAACGAGCTGTTCGATC 60  
 QY 204 CGCCGCGACGGCGATCCCGGCGCGAAGCCGCCGCGGAGCAGAACGGCGAGTGTGCCAG 263  
 Db 61 CGACATGAATCTGATGCGCGGACCAACCTTTGAAGATATGAATGGCCAGTATGTCAG 120  
 QY 264 ATTGCGCGCAGCAGCGTCGCGCTTGCCCCCGCGCGGAGACCCCTTCTGTGCTGCACGAG 323  
 Db 121 ATCTGTGTGATGATGATGTTGGACTCGCTGAACCTGGAGATGCTTTGTCCGCTGTAATGAA 180  
 QY 324 TGC GCTTCCCCGCTCTGCGCGGACTGTCTACGAATACGAGCGCGGAGAGGCACGCGAGAAC 383  
 Db 181 TGTGCTTCCCTGTGTGTGCGCCCTTGCTATGAGTACGAGAGGAAGATGAACTCAGTGT 240  
 QY 384 TGCCCCCAGTGCAGACTCGATACAGCGCCTCAAGGGCTGCCAACGTGTGACCGGTGAC 443  
 Db 241 TGCCCTCAATGCAAGACTAGATTACAGACGCACAGGGGGAGTCTCTGTGTGAAGAGAT 300

QY	444	GAGGAGGAGACGGCGCTGCATGACCTGGACCAACGAGTTCAACTGGAGCGGCATGACTCG	503
Db	301	GAAATGAGATGATGATGATGATATCGAAGATGAGTTCA-----ATTAGGCC	348
QY	504	CAGTCTGTGCGGAGTCCATGCTCTACGGCCACATGAGCTACGGCCGTGAGGTGACCT	563
Db	349	CAGGAGCTAACCAAGCGAGACACCAACGCCA---TGGGAGAAGAGTTTCTTCTTCTCT	405
QY	564	AATGGCGGCCCAAGCTTTCAGCTCAACCCCAATGTTCACCTCCTCACCAACGGGCAA	623
Db	406	AGACATGAATCTCAACCAATTCTCTTCTCACCCCATAGCCATACGGTTCTTGAGAGATT	465
QY	624	ATGTTGATGACATCCACCGGAGCAGCAGCGCGTGTGCTTCTTCATGGTGTGG	683
Db	466	CGACGCGCTGATACCAATCTGTGCGAACTACATCAGTCTT-----TGGTCTCT	519
QY	684	GGAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGTCT	743
Db	520	GACAGAGATGCTATTCTCATCTCATATATTGATCCAGGCAACCTGTCCCTGTAAGATC	579
QY	744	ATGGAACCATCCAGAATCTTGCTGCATATGGGTATGTGTGCTTGTGAAGAACGG	803
Db	580	GTGACCCCGTCAAAAGACTTGAACCTTATGGGCTTGTAATGTTGACTGAAAGAAAGA	639
QY	804	ATGGAATTTGAGACGAGACAAGAGAGATGC-----	837
Db	640	GTTGAAGCTGGAAGCTGAGACAGAGAAAAATATGTACAGATGACTGTAATACCAT	699
QY	838	-ACCAGACGGGAATGATGTGTGTGATGATGTGACGATGCTGATCTACCACTAATG	896
Db	700	GAAAGGAAAGAGAGAAATTGAAGGAGACTGTTCCAATGGCGAAGAACTCCAAATGGCT	759
QY	897	GATGAAGCAAGACAACACTGTCCAGAAATTCACCTTCCATCAAGCCAGATTATCCA	956
Db	760	GATGATACACGCTCTTCTATGAGTCGTGTGTCCTATCCCATCTTCTCGCCTAACCCCT	819
QY	957	TATAGATGATTATCATTAATCCGCTGTGTGTTTGGGATTTCTTCCACTACCGAGTG	1016
Db	820	TATCGGCTGTGATTATTTCCGGCTTATCATCTTGTTCTTCTTGCAATATCGTACA	879
QY	1017	ATGCATCCGGTGAATGATGATTTGCTTGTGTGCTCATCTGTTATCTGTGAATCTGG	1076
Db	880	ACTCACCCGTGAAAAATGCATATCCTTGTGTGTAAGCTCGGTATCTGTGAGATCTGG	939
QY	1077	TTTGCCATGTCTTGATCTTGATCAATTCGCCAAAGTGTCCCTATGAGAGAGAGACT	1136
Db	940	TTTGCAATTTCTTGGCTCTTGATCAGTTTCCCAATGGTACCCCATTAACAGGAGACT	999
QY	1137	TACCTAGACCGGCTGTCACTGAGGTTGCACAAGGAAGGCCAGCCATCTCACTTGTCCA	1196
Db	1000	TATCTTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAAACCATCACAGCTCGTCT	1059
QY	1197	ATTGATTTCTTTGTCAGTACGGTTGATCCCTTAAAGAACCTCCTTGTGTCAACAACAT	1256
Db	1060	GTTGATGTGTTGTAGTACAGTGAGACCCCATTGAAAGAGCCTCCCTGTGTACAGCAAC	1119
QY	1257	ACTGTTCTATCTATCTTTCGGTGATTAATCTGTGTAAAGGTTTCTTGCTATGTTTCT	1316
Db	1120	ACAGTCTCTCGATTCTTCTGTGACTACCCGGTAAAGTAAAGTAGCCTGTATGTTTCA	1179
QY	1317	GATGATGTCGTCAATGCTAACGTTTGAAGCATTAATCTGAACAATCTGAATTTGCAAG	1376
Db	1180	GATGATGTCGTCAATGCTTAACCTTGAATCCCTTCTGAACAACGCTGAGTTGCAAG	1239
QY	1377	AAATGGTTCTTCTGCAAAACGCTACAAATATTGAACCTCGCGCTCCAGAGTGTAATTC	1436
Db	1240	AAATGGGTACCATTTTGCAAGAATTTCAACATTGAACCTAGGGCCCCCTGAATTTCTAATTT	1299
QY	1437	CAACGAAGATAGACTTAATGAAAGACAAGGTGGCAGCAAACTTTGTTAGGGAGAGAGGA	1496
Db	1300	GCCCAAGATAGATTACTTGAAGGACAAGATCCACCGTCTTTGTTAAAGGAGCAGGA	1359

Accession	Sequence	Position
Db	1060 GTTGATGTGTTGTTAGTACAGTGGACCCATTGAAAGACCTCCCTTGTTACAGCAAAAC	11199
QY	1257 ACTGTTCTATCTATCCCTTCGGTGATTAATCCGTGTGATAAGTTTCTTGCTATGTTTCT	13166
Db	1120 ACAGTTCTCTCGATTCTTTCTGTGACTACCCCGGTAGATPAAGTAGCCTGTTATGTTTCA	11799
QY	1317 GATGATGGTGTGCAATGCTAACGTTGGAAGCATTAATCTGAAACATCTGAATTTGCCAAG	13766
Db	1180 GATGATGGTTCAGCTATGCTTAACCTTGAATCCCTTTCTGAAACCGCTGAAGTTGCCAAG	12399
QY	1377 AAATGGGTTCCCTTTCGCAAAACGGTACAATATTGAACCTCCGCTCCAGAGTGGTACTTC	14366
Db	1240 AAATGGGTACCATTTTGCAAGAAATTCAACATTGAACCTTAGGGCCCCCTGAATTTCTAATTT	12999
QY	1437 CAACAGAGATAGACTACTTGAAAGACAAAGGTGGCAGCAAACTTTGTTAGGAGAGAGAGA	14966
Db	1300 GCCCAGAGATAGATTACTTGAAGGACCAAGATCCAAACCGTCTTTGTTAAAGGCGCAGA	13599





; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 3614  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (217)..(3411)  
US-09-221-013A-9

Query Match 32.2%; Score 1269.2; DB 4; Length 3614;  
Best Local Similarity 63.8%; Pred. No. 0;  
Matches 2051; Conservative 0; Mismatches 1108; Indels 57; Gaps 6;

```
QY 209 CGACGGCATCCCGGGCGGAGCCGCGGAGACAGAACGGGCAAGTGTGCCAGATTG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 CGAAGGAGAAACCGCGGAAAGCCGATGAAGAACATTGTTCCGCAGACTTGCCAGATCTG 284

QY 269 CGGCGACGACGTGGCCTTGCCCCCGGGGGGACCCCTTCGTGGCGTGCAACGAGTCCG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 TAGTGAACAATGTGGCAAGACTGTTGATGAGATCGTTTGTGGCTTGATATTGTTTC 344

QY 329 CTTCCCCGTCTGCGGGACTGTGCAATAAGAGCGCCGGAGGCAACGAACTGCCCC 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 ATTCGCAAGTTGTGCGCCTTGCTACGAGTATGAGAGAAAGATGGAAATCAATCTGTCC 404

QY 389 CCAGTGCAAGACTCGATACAAGCGCTCAAGGGCTGCCAACGTGTGACCGGTGACAGGA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 TCAGTGCAAAACAGATACAGAGGCTCAAGGATAGTCTGCTATCTCTGTGATTAAGA 464

QY 449 GGAGGACGGCGTCGATGACCTGGACACAGAGTTCACACTGGAGCGGCATGACTCGCAGTC 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 CGAGGATGGCTTAAGCTGATGAAGTACTG--TTGAGTTCAACTACCTTCAGAAAGAGAA 521

QY 509 TGTGGCCGATCCATGCTCTACCGCCACATGAGCTACGCGCGTGAGAGTGAACCTAATGG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 522 AATTTCAGAGCGGATGCTTGGTTGGCACTTACTCTGTTGGAGAGAGGAAATGGGGGA 581

QY 569 CGGCGCACAGCTTTCAGCTCAACCCCA--ATGTTCCACTCTCTCAACCAAGGCAAT 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 ACCCCAGTATGATAAAGAGTCTCTCAATCATCTCTCTCTCAACGACAGACAAGA 641

QY 626 GGTGATGATCCACCGAGACGACGCGCTGTGCTTTCATGGGTGTGGGG 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 TACTTCAGAGAGATTTTCTGCTGCTCACTGAACGCTCTCTGATCTTCTACTATCGC 701

QY 686 AAAGAGATATACCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGGTCTAT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 702 TGGGGGAAACGCCCTTCCCTATTCTATCATGATGTAATCAATCAACCAATAGAAAGATTGT 761

QY 746 GGACCATCCAAAGATCTTGTGCATATGGGTATGTAAGTTGCTTGGAAAGAACGAT 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 762 GGATCTGTGGAATCGGGAATGAGTTGAAAGAGAGAGATGAGCTGGAAAATGAA 821

QY 806 GGAGAAATGGAAGCAGAGACAAGAGAGATGACACGAGGGAATGATGTTGTGTA 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 822 GCAAGAGAGAAATACTGCTCTGTACAGACGAGGCTGCTTGAAGAGAGTGAAGTAGA 881

QY 866 TGATGTGACGATGCTGATCTACCA-----CTAATGATGAAGCAAGACA 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 882 TATTGATGCCACACAGATATCTAGCAGATGAGGCTCTGCTGAATGACGAAGCGAGCA 941

QY 911 ACAACTGTCCAGAAATATCCACTTCCATCAAGCCAGATTATCCATATAGATGATTAT 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 942 GCTTCTGTCAAGAAAGTTCAATTCCTTCATCAAGGATCAATCTTACAGAAATGTTAT 1001

QY 971 CATTATTGGCTGTGTTTGGGGTCTTCTTCCACTACCGAGTGAATCCGGTGAA 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1002 TATGCTGCGGCTGTATCTTGTCTCTTCTGCAATTACCGTATTAACAAACCAAGTGC 1061

QY 1031 TGATGATTTGCTTGTGCTCATATCTGTATCTGAAATCTGTTGCGCATGTCTTG 1090
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1062 AAATGCCCTTGTCTATAGCTGGTCTCTGTGATATGAGATCTGTTGCCATTATCTG 1121

QY 1091 GATTCCTGATCAATTCCAAAGTGTTCCCTATTTGAGAGAGACTTACCTAGACGGCT 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1122 GATTTTGATCAGTTTCCCAAGTGTTTCCGTGAAACCGTGAACCTTACCTCGACAGCT 1181

QY 1151 GTCACTGAGTTCGACAGAGAGCCAGCATCTGCACTGTCCCAATGATTTCTTGT 1210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1182 TGCTTTAAGATATGATCGTGAAGGTGAGCCATCAAGTTAGCTGTGATCATTTGCT 1241

QY 1211 CAGTAACGTTGATCCCTTAAAGAACTCTTGTGTTGTTCAACAACAATACTGTTCTAT 1270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1242 GAGTACTGTTGACCCCTTGAAGAGACCCACCCCTTGTGACAGCCACACAGTGTCTAT 1301

QY 1271 CCTTTCGGTGAATTATCTGTGATTAAGTTCTTGTATGTTCTGATGATGTGCTGC 1330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1302 TGTGGCTGTGACTACCAAGTTGACAAAGGTGCTGTATGTTCTGATGATGGTGTGC 1361

QY 1331 AATGCTAACGTTGAAGCATTAATCTGAACATCTGAATTGCAAAAGAAATGGTCTTT 1390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1362 TAGTTATCATTTGAATCACTTGCAAGAAACATCAGAGTTTGTCTGTAATGGTACATT 1421

QY 1391 CTGCAAAACGTTACAATATGAACCTCGCGCTCCAGAGTGTACTTCCAACAGAGATAGA 1450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1422 TTGCAAGAATATAGCATAGACCTCGTGACACAGAAATGTAATTGCTGCCAAATAGA 1481

QY 1451 CTACTTGAAGACAAAGGTGGACCAAACTTGTGTAGGAGAGAGAGCAATGAAGAGA 1510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1482 TTACTTGAAGATTAAGTTGACAGATCATTTGTCAAGATCTGATGATGAAGAGGA 1541

QY 1511 GTATGAGGAATTCAGGTGGAATCAATGCTTGTAGTCCAAAGCCCAAGAAATCTGTA 1570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1542 ATATGAGGAATTTAAATCCGAATCAATGCACTGTTTCCAAAGCCCTAAATGTCTGA 1601

QY 1571 AGAAGATGACAATGCAAGATGAAGAAACCCCTGGCTGGAACAATGTTGATGATCC 1630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1602 AGAAGGTGGGTATGCAAGATGGCACACCGTGGCTGGAATAATACAGGGACCATCC 1661

QY 1631 TGAATGATTCAGTCTTCTTGGCCAAAGCGAGGCTTGACTGTGAGGAATGAACT 1690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1662 AGGAATGATCCAGGTCTTCTTAAGGCAAAATGTTGATGATGACAGAGGCAATGACT 1721

QY 1691 GCCAGATTGTTATGTTCTAGAGAAACGACCAAGCTATAACCATCATTAAGAAAGC 1750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1722 CCGCGGTTGGTATATGTTCTCTGAGAAAGCGACAGGATTCACGACCACAAAAGGC 1781

QY 1751 TGTGCTATGAATGATGCTTGGCTGCTGTGTAATAACAATGCTCCATATTTGTT 1810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1782 TGGTGTATGAATGACATGCTGAGAGATTTCAGCAGTTCTTAACAATGACCTTTCATCT 1841

QY 1811 AAATCTGATTGTGATCACTACATCAACAACAGCAAGCTATAAAGGAAGCAATGTGTT 1870
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1842 GAATCTGATTGTGATCATATATAAATACAGCAAGCTTAAGAGAACATGTGCTT 1901

QY 1871 TATGATGACCCCTTACTAGAAAGAGTTTGTCTATGTACAGTTCCTCAAGATTGTA 1930
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1902 CCTGATGACCCCAACCTCGGGAAGCAAGTTGTTATGTTCAAGTTCACCAAGATTGTA 1961

QY 1931 TGGGATGATCGCCATGACCATATGCTAACCGGAATGTTCTTTTGTATATCAACAT 1990
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1962 TGGTATCGATTAAGAACGATAGATATGCTAATCGTAATACCGTGTCTTGTATATTACTT 2021

QY 1991 GAAAGTTTGGATGTTATCAGGCTCAATTTATGTTGTAAGTGAATGTAATTAGAG 2050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2022 GAGAGTTTGAATGAGATTCAGAACTGTATATGTCGAACCTGATGTTTCAACAG 2081

QY 2051 GCAGCATTATATGTTATGATGCCCCCAAAACAAAGAACCCACATCAAGACTTGCAA 2110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2082 AACAGCATTATACGTTATGAACCTCCAATAAAGTAA----- 2120

QY 2111 CTGCTGCCCAAGTGTGCTTTGCTGTGCTGCTTGGCAATAGAGAACAAAGAAC 2170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2121 -----ACACAAGAGCCAGTCTTTATCTAAGCTCTGTGTGATCAAGAAAGAGAA 2174
```

```
QY 2171 TACCAACCCAAAACAGAGAAAAAGTTATTTTTCAGAGAGAGAACCAATC 2230
    |||||
Db 2175 TTCCAAAGCTAAGAAAGAGTCGGACAAAAAGAAATCAGGCAGGACATGACTCAACTGT 2234
QY 2231 CCCTGCATATGCTCTTGGTGAATTCAGCAAGCTGCTCCAGAGC-----TGAGATGA 2284
    |||||
Db 2235 TCCTGTATCAACCTCGATGACATAGAGAGGAGGTGAAGGTGCTGGTTTGTATGATGA 2294
QY 2285 AAAGCCGGTATGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCAACTTCTGTTTT 2344
    |||||
Db 2295 AAAGCGCTCTTAATGTCCGAAATGAGCCTGAGAAAGCATTTTGACACAGTCTGCTGTTTT 2354
QY 2345 TGTACATCCACACTTCTCGAGATGTGGAACCTTGAAGAGTGCAAGTCTGCTTCTCT 2404
    |||||
Db 2355 TGTGCTTCTACCTTAATGGAATGAGTGTGTTCTCTTCCAGCAACTCCAGAAAACTT 2414
QY 2405 TTTGAAAGAGCTATACATGTCATTAGTTGTGTTATGAAGACAGACAGACTGGGAAA 2464
    |||||
Db 2415 TCTCAAAGAGCTATCCATGTCATTAGTTGTGTTATGAGGATAAGTCAGATTGGGGAAT 2474
QY 2465 AGAGATTGGCTGATCTATGATCAGTTACAGAAAGATATTCTAAGTCTGTTCAAGATGA 2524
    |||||
Db 2475 GGAGATTGATGATCTATGTTCTGTGACAGAGATATTCTGAGTGGGTTCAAAATGCA 2534
QY 2525 TTGTCATGTTGGCGGTCAATTACTGCATACCTAAACGGGTTGCATTCAAGGTTCTGC 2584
    |||||
Db 2535 TGCCCGTGATGGCGATCCATTACTGATGCTTAAGCTTCCAGCTTCAAGGGTCTGC 2594
QY 2585 ACCTCTGAATCTTTCAGATCGTCTTCCACAGGTCTCGGTGGCTCTTGGGTTATTTGA 2644
    |||||
Db 2595 TCCTATCAATCTTTCAGATCGTCTGAAACCAAGTCTGAGTGGGCTTACAGTTCAAGTTGA 2654
QY 2645 GATCTTCTCAGCAATCATGCCCCCTCTTGGTATGGGTATGGTGGCGGTGAATTTTTT 2704
    |||||
Db 2655 GATTCTCTCAGTCGGCATTTCTCTATATGTTATGTTTACAATGGAGGCTAATAATTTCT 2714
QY 2705 GGAAGATTTTCTTACATCACTCCATCGTATCTTGGACATCTATTCCTCTTGGC 2764
    |||||
Db 2715 TGAGAGTTTGCGATGTGAACACCAACCATCTACCTCCATCTCTCTCAT 2774
QY 2765 TTACTGTACATGCTGCCATCTGTTTATGACAGGAAATTTATCCTCAGAGCTGAA 2824
    |||||
Db 2775 GTATTGTACATGCTAGCCGTTTGTCTCTTCCACCAACAGTTTATTTCTCAGATTAG 2834
QY 2825 TAATGTTGCCAGCCTGCTCATGTCATTTTATCTGCATTTTGTCTACAGCATCCT 2884
    |||||
Db 2835 TAACATTGCAAGTATATGTTTCTGTCTCTCTCTCTCTCCATTTTCCGCGGATTAAT 2894
QY 2885 AGAATGAGATGAGTGTGTTGGAATGATGATGATGAGGAAATGAGCACTTCTGGGT 2944
    |||||
Db 2895 AGAATGAGTGTGAGTGTGAGGATGAGCAATGATGAGGAAACGAGCAGTTTGGGT 2954
QY 2945 CATTGAGGTGTGCTCACAACCTCTTGTGTGTTCAGGGACTTCTCAAGTCAATAGC 3004
    |||||
Db 2955 CATTGTGAGTATCGCTCATTTATTGCTGTGTTCAGAGTATCTCAAGTCTTGC 3014
QY 3005 TGGTGTGATACAAGCTTCAACCGTACATCAAGGAGGAGATGATGA--GAGTTCTC 3061
    |||||
Db 3015 CGGTATGACACAACCTTCAAGTACCTCAAAAGCTTCAAGTGAAGAGGAGACTTTCG 3074
QY 3062 AGAGCTATATACATCAATGAGTACCTTATGATACCTCCTACACCTGCTTCTAT 3121
    |||||
Db 3075 TGAGCTCTACTTGTGAATGAGACAACACTTCTGATTCGCGCAACGAGCTGCTCATGT 3134
QY 3122 GAACCTCATTTGTTGCTGCTGCGCTTCAAAATGCGATCAATAACGATATGAGTCATG 3181
    |||||
Db 3135 AAACCTTAGTGGAGTTGTGACAGAGTCTCTTATGCTATCAACAGTGATACCAATCATG 3194
QY 3182 GGGCCCCCTCTTGGGAAGCTATTTTGCATTTTGGGTGATTTGCCATCTTTATCCCTT 3241
    |||||
Db 3195 GGGACCACTCTTGTGTAAGTTGTCTTGTGCTTCTGCGTGAATGTTTCACTTGTACCTTT 3254
```

```
QY 3242 TCTCAAGGTTTGGTGAAGGCAAAACAGACACCAAGATGTCATGCTGTGTCAT 3301
    |||||
Db 3255 CCTCAAGGTTTGTATGGTGCACAGAACCGGACTCTTACCAATGTTGTGTGTGTGTGT 3314
QY 3302 TCTGCTGCTTCAATCTTCTGCTCTTGGGTTGGGTTGGATGATCTTCTTGGCAAGA 3361
    |||||
Db 3315 TCTCTGGCTTCTATCTTCTGTTGTGTGGTTAGATGATTCCTTCACTAGCCGAGT 3374
QY 3362 TGATGTCGCTTCTTGAGAGTGTGTTGATTG 3397
    |||||
Db 3375 CACTGGCCCGGACATTTCTGGAATGTGAATCAACTG 3410
```

## RESULT 5

```
US-08-960-048-1
; Sequence 1, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-1
```

```
Query Match 28.8%; Score 1134.8; DB 3; Length 3328;
Best Local Similarity 66.1%; Pred. No. 5.8e-312;
Matches 1731; Conservative 0; Mismatches 827; Indels 60; Gaps 4;

QY 768 GCATATGGGTATGTTAGTGTGCTTGGCAAGAACGGATGAGAAATTGGAAGACAGACAA 827
    |||||
Db 488 GCTGAAGACAATGGGAATTCGATTGGAAAGAACAGGTTGAAAGTGAAGAAAGAAAGAG 547
QY 828 GAGAGATGACACGAGCGGGAATGATGTTGTGTGATGATGATGATGATGATGATGATG 887
    |||||
Db 548 AACAAAGAAAGAACCTGCAACAACTAAGTTGAAAGAGAGGCTGAAATCCACCTGAG 607
QY 888 CCACTAATGA-----TGAAGCAAGACAACTGTCCAGAAAAATTCGA 932
    |||||
Db 608 CAACAAATGGAAGATAAACCGGACCGGATGCTTCCAGCCCTCTGACTATAATTCCA 667
QY 933 CTTCCATCAAGCCAGATTAATCCATATAGATGATTATCATTAATTCGCTTGTGTTTG 992
    |||||
Db 668 ATCCGAAAGACAGACTTGCACCAATACGAAACCGTATCATTAATGATGATCATTTCT 727
QY 993 GGGTCTTCTTCCACTACCGAGTGCATCCGGTGAATGATGATGATGATGATGATGATG 1052
    |||||
Db 728 GGTCTTTTCTTCCATTATCGAGTACAAACCCCGTTGACAGTGTCTTTGACTGTGCTC 787
QY 1053 ATATCTGTATCTGTGAATCTGGTTTGCATGTCTTGATTTCTTGATCAATTCCCAAAG 1112
    |||||
Db 788 ACTTCAGTATATGTGAATCTGGTTTGCATTTTCTGCGGTGTGATCAAGTCCCTAAG 847
QY 1113 TGGTTCCTATTGAGAGAGAGACTTAAGTGAACCGGCTGTCACTGAGGTTGACAAGGAA 1172
    |||||
Db 848 TGTATCTGTAAACAGGAAACATACATGACAGACTATCTGCAAGATATGAAAGAGAA 907
QY 1173 GGCCAGCAATCTCAACTGTCTCAATTGATTTCTTGTGATGATGATGATGATGATGATG 1232
    |||||
Db 908 GGTGAACCTGATGAACCTTGTGAGTGTGATCTTCTGTGATGATGATGATGATGATGAA 967
QY 1233 GAACCTCTTGTGTCACAAACAATCTGTCTATCTATCTCTTCCGTTGGTGAATATCTGT 1292
```

Db 968 GAGCTCCATTGATTACTGCGCAATACTGTGCTTCCATCCCTGCTTGAGTACCCGGTG 1027  
Qy 1293 GATAAGGTTCTTGCTATGTTTCTGATGATGCTGCAATGCTAAAGTTGAAGCATTA 1352  
Db 1028 GATAAGGTTCTTGTTATATATCTGATGATGCTGCGGCATGCTGACATTTGAATCTCTA 1087  
Qy 1353 TCTGAAACATCTGAATTTGCAAGAAATGGGTTCTTCTGCAAAACGGTACAATATTGA 1412  
Db 1088 GTAGAAACAGCCGACTTTGCAAGAAAGTGGGTTCCATTCTGCAAAAATTTTCCATTGAA 1147  
Qy 1413 CCTGCGGCTCCAGAGTGTACTTCCACAGAGATAGACTACTTGAAGAACAAGTGGCA 1472  
Db 1148 CCGCGGCACTGAGTTTACTTCTCACAGAAAGATTGATTACTTGAAGATAAAGTGCAG 1207  
Qy 1473 GCAAACTTTGTTAGGAGAGAGACAATGAAGAGAGATATGAGAAATTCAGGTGAGA 1532  
Db 1208 CCTCTTTGTAAAGAACGTAGAGCTATGAAAAGAGATTATGAAAGTACAAAATTCGA 1267  
Qy 1533 ATCAATGCTTAGTTGCCAAAGCCGAAAGTTCTCTGAAGAAGATGACAATGCAAGAT 1592  
Db 1268 ATCAATGCTTAGTTAGTTCAAAGGCTCAGAAAACACCTGATGAAGATGACATGCAAGAT 1327  
Qy 1593 GGAACCCCTGGCCTGGAACAATGTTGCTGATCATCCTGGAATGATTCAGGCTTCTT 1652  
Db 1328 GGAATTTCTTGCCAGAAATAACCCGCGTGATCACCTGCGCATGATTCAGGTTTCTT 1387  
Qy 1653 GGGCAAAAGGAGGCGCTTGACTGTGAGGGAATGAACCTGCCACGATGGTTATGTTCT 1712  
Db 1388 GGATATAGTGTGCTCGTGACATCGAAGGAATGAACCTTCTGCACTGGTTTACGTTCT 1447  
Qy 1713 AGAGAGAAACGACCGACTATTAACCATATCAATAAGAAAGCTGTGCTATGAATGATGATC 1772  
Db 1448 AGAGAGAAAGACCTGCTTACCAACACCAAAAAGCGTGGTGAATAATGCTTGGTT 1507  
Qy 1773 CGAGTCTCTGCTGTAATAACAATGCTCCATATTGTTAACTTGATGATGATCATCTAC 1832  
Db 1508 AGGTTGCTGCAAGTTCTTACAAATGCTCCCTCATCCTCAATCTTGAATTGTGACCACTAT 1567  
Qy 1833 ATCAACAACAGCAAGGCTATAAGGAACAATGTGTTTATGATGGAACCCCTTACTAGGA 1892  
Db 1568 GTTAACAATAGCAAGGCAAGGATGAGGCAATGTGCTTGTGATGGAACCCACAAGTTGTT 1627  
Qy 1893 AAGAAAGTTTCTATGTACAGTTCCCTCAAAAGATTGATGGAATGATGCCATGACCGA 1952  
Db 1628 CGAGATGATGCTATGTGAGTTCCCTCAAAAGATTGATGGAATGATGAGATGATCGA 1687  
Qy 1953 TATGCTAACCGAATGTGCTTTTGTATATCAACATGAAGGTTGGATGATTCAG 2012  
Db 1688 TATGCAATAAGGAACACAGTTTCTTGTATGTTAACAATGAAGGCTTGATGAATCCAA 1747  
Qy 2013 GGTCAATTATGTTGTAAGTGTATTTAGAAAGGCAAGCATTATATGTTATGAT 2072  
Db 1748 GGGCCAGTTATGTGGAACAGGTTGTTTCAATAGGCAAGCACTTATGGCTATGGT 1807  
Qy 2073 GCGCCCAAAACAAAGAGCCACCATCAAGACTTGCAACTGCTGGCCCAAGTGGTCTT 2132  
Db 1808 CCACCTTCAATGCCAA-----GTTTCCCAAGTCATCTCC 1843  
Qy 2133 TGCTGTGCTGCTTTGGCAATAGAGCAAAAAGAGACTACCAAAACCAAAACAGAGAA 2192  
Db 1844 TCATCTTGCTGCTGTGCTGCGCGGCAAGAAAGAACCTTAAGATCC----- 1890  
Qy 2193 AAAAAGTATATTTTTCAGAAAGAGAAACCAATCCCTGCATATGCTCTGTGTA 2252  
Db 1891 -----ATCAGAGCTTATAGGATGCAAAACGGGAAGAACTTGATGCTGCCATCTTAAC 1945  
Qy 2253 ATTGACGAAGCTGCTCCAGAGCTGAGATGAAAAGCCGGTATTGTAATCAACAAAA 2312  
Db 1946 CTTAGGGAATTGACAATTATGATGATGAAAAGATCAATGTGATCTCTCAACAAAGC 2005  
Qy 2313 TTAGAAAAGAAATTTGGCCAATCTTCTGTTTGTGTACATCCACACTTCTCGAAGATGGT 2372

Db 2006 TTGAGAAACTTTGGCTTATCTTCAGTCTTCAATTGAATCTACACTAATGAGAAATGGA 2065  
Qy 2373 GGAACCTTGAAAGATGCAAGTCTGCTTCTCTTTGAAAGAAAGCTATATACATGCTATTAGT 2432  
Db 2066 GAGATGGCTGAATCTGCAACCCCTTCCACACTAATCAAGAAAGCAATTCATGCTCAAGC 2125  
Qy 2433 TGCTGTATGAAGACAAGACAGACTGGGAAAAAGATTGGCTGATCTATGATGATCAAGTT 2492  
Db 2126 TGTGCTATGAAGAGAGACTGCATGGGGAAGAGATTGGATGATATATGTTCAAGTC 2185  
Qy 2493 ACAGAGATATTTCACTGCTTCAAGATGATGTCATGTTGGCGTCAATTTACTGC 2552  
Db 2186 ACTGAGATATCTTAACCGGCTTCAAAATGCATGCCAGAGATGAGATCGATTTACTGC 2245  
Qy 2553 ATACCTAAACGGGTTGCATTTCAAAGGTTCTGCACCTCTGAATCTTCAAGATCGTCTTCA 2612  
Db 2246 ATGCCCTTAAGCCAGCATTCAAAGGATCTGCACCCATCAATCTGTGATCGGTTGCAC 2305  
Qy 2613 CAGGTCTTCGGTGGCTCTTGGGCTATTGAGATCTTCTCAGCAATCATGCTCTT 2672  
Db 2306 CAGGTCTTCGATGGCTCTTGGATCTGTTGAATTTTCTTAAGCAGGCAATGCCCTTA 2365  
Qy 2673 TGGTATGGGTAT--GGTGGCGTCTGAAATTTTGAAGAAATTTTCTACATCAACTCC 2729  
Db 2366 TGGTATGGCTTTGAGGTGGTGGTCTTAATGGCTTCAAGAACTAGCATATATAACACC 2425  
Qy 2730 ATGCTGATCCTTGACATCTATTCGCCCTCTTGGCTTACTGTACATTCCTGCCATCTGT 2789  
Db 2426 ATTGCTATCCTTTCACATCCCTTCCACTCATTCGCTAATGTTCACTACCAAGCAATCTGT 2485  
Qy 2790 TTATTGACAGGAAATTATCACTCCAGAGCTGAAATATGTTGCCAGCCTGTGTTCATG 2849  
Db 2486 CTTCTCACAGAAATTTATCATACCAACGCTTCAAAACCTGGCAAGTGTCTTCTTCTT 2545  
Qy 2850 TCACTTTTATCTGCAATTTTGTCTACGAGCATCTGAAGATGAGATGAGTGGTGTGA 2909  
Db 2546 GGCCTTTCTTCCATTAATGCTGACTGTGTTCTGAGCTCCGATGAGTGGTGTCAAGC 2605  
Qy 2910 ATTGATGACTGTGAGGAATGAGCAGTTCTGGTCAATTGGAAGTGTCTCTCAACCTC 2969  
Db 2606 ATTGAGGACTTATGGCGTAACGAGCAGTTTGGTTCATGGTGGCGTTCAAGCCATCTC 2665  
Qy 2970 TTTGCTGTGTTCCAGGACTTCTCAAGGTCATAGCTGTGTTGATACACAGCTTCAAGCTG 3029  
Db 2666 TTTGCCGCTTCCAAAGTTTCTTAAGATGCTTGGCGGCAATTGACACCACTTTACTGTG 2725  
Qy 3030 ACATCAAAAGGTGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAAATGACTAAC 3089  
Db 2726 ACTGCCAAAGCAGCTGATGATGCAAGATTTGGTGAAGCTCTACATTTGAAATGACTACA 2785  
Qy 3090 TTATGTATACCTCTACCACTTGTCTTATTTGAATTCATTTGGTGTGCTGGCGTT 3149  
Db 2786 CTTCTAATCCCTCAACAAACACTCCTCATGCTCAACATGTTGGTGTGCTGGCGAATTC 2845  
Qy 3150 TCAATGCGATCAATAACGATATGAGTCATGGGCCCCCTTTTGGAAAGCTAATCTTT 3209  
Db 2846 TCCGATGCCCTCAACAAAGGTAACGAGCTTGGGACCACTTTTGGCAAGTGTCTTT 2905  
Qy 3210 GCATTTTGGGTGATTTTCATCTTTATCCCTTTCTCAAAAGGTTTGGTGAAGGCAAAAC 3269  
Db 2906 TCTTCTGGGTCACTCTTCATCTTTATCCATTTCTCAAAAGGTTTATGGAGCGCAAAAC 2965  
Qy 3270 AGGACACCAACGATTGTCATGCTGTGTCATTTCTGTGGCTTCAATCTTCTGCTCTT 3329  
Db 2966 AGGACACCAACCAATTTGTTGCTTGTGTCAGTGTGTTGGCTTCTGTCTTCTTGT 3025  
Qy 3330 TGGGTTCGATTGATCTTCTCTTGGAGAGATGATGG 3367  
Db 3026 TGGGTTCGATCAACCCGTTTGTCAACACCGCGATAG 3063

RESULT 6  
US-09-838-586-1



; Sequence 1, Application US/09838586  
; Patent No. 6576818  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE OF INVENTION: Sequences  
; FILE REFERENCE: 15621/02/US  
; CURRENT APPLICATION NUMBER: US/09/838,586  
; CURRENT FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/029,987  
; PRIOR FILING DATE: 1996-10-29  
; PRIOR APPLICATION NUMBER: 08/960,048  
; PRIOR FILING DATE: 1997-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3328  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Oligonucleotide  
US-09-838-586-1

Query Match 28.8%; Score 1134.8; DB 4; Length 3328;  
Best Local Similarity 66.1%; Pred. No. 5.8e-312;  
Matches 1731; Conservative 0; Mismatches 827; Indels 60; Gaps 4;

QY 768 GCATATGGGTATGGTATGTTGCTTGGAGAGAACGGATGGAGATTGGAACAGAGACAA 827  
DB 488 GCTGAAGACATGGGAATTGCAATTTGGAGAGAACGGGTGGAAGTTGGAAGAAAGAA 547  
QY 828 GAGAGATGCACGACGGGGAATGATGGTGGTGATGATGGTGACATGCTGATCTA 887  
DB 548 AACAGAGAGAGAGAGCTGCAACACTAAGTTGAAAGAGAGAGGCTGAAATCCCACTGAG 607  
QY 888 CCACCTAATGA-----TGAAGCAAGACAACAACCTGTCAGGAAATTCCA 932  
DB 608 CAACAATGAGAGATAAACCGGACCGGATGCTCCAGCCCTCTGACTAATTAATTCCA 667  
QY 933 CTTCCATCAAGCCAGATTAATCCATATAGATGATTAATCATTAATCGGCTTGTGTTTG 992  
DB 668 ATCCGAAAGAGAGAGCTGCAACCATACCGAACGGTATCATTAATGCAATGATCATTTCT 727  
QY 993 GGGTCTTCTTCCACTACCGAGTGTATGATCCGGTGAATGATGATGATTTGTTGGCTC 1052  
DB 728 GGTCTTTTCTTCCATTAATGAGTAAACAAACCCCGTTGACAGTGTCTTTGGACTGTGCTC 787  
QY 1053 ATATCTGTTATCTGGAATCTGTTGCCATGCTTGGATTCTTGATCAATCCCAAG 1112  
DB 788 ACTTCAGTCATATGTAATCTGTTGTCATTTCTCGGTGTTGATCAATTCCTTAAG 847  
QY 1113 TGGTCCCTATGAGAGAGAGACTTACCTAGACCGGCTGTCACTGAGGTTGACAGAA 1172  
DB 848 TGGTATCCTGTTAACAGGAAACATACATTGACAGACTATCTGCAAGATATGAAAGAA 907  
QY 1173 GGCCAGCCATCTCAACTTGTCCCAATTGATTTCTTGTGACAGTACGTTGATCCCTTAAG 1232  
DB 908 GGTGAACCTGATGAACTTGTGAGTTGACTTCTCGTAGTACAGTGAATCCATTGAAA 967  
QY 1233 GAACCTCCTTGTGCAACAACAATACTGTTCTATCTATCCCTTCCGTTGATTAATCCTGTT 1292  
DB 968 GAGCCTCCATTTGATTAATGCAATCTGTGCTTTCATCTTGCCTTGAATCCCGGTG 1027  
QY 1293 GATAAGGTTTCTGCTATGTTTCTGATGATGGTGTGCAATGCTAACGTTGAAGCATTA 1352  
DB 1028 GATAAGGTTCTTGTATATATCTGATGATGGTGGCCCATGCTGACATTTGAATCTCTA 1087  
QY 1353 TCTGAACATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACCGGTACATATTGAA 1412  
DB 1088 GTAGAAACAGCCGACTTTGCAAGAAAGTGGGTTCCATTCTGCAAAAAAATTTTCCATTGAA 1147  
QY 1413 CTTGGCGCTCAGAGTGTACTTCCAAAGAGATAGACTTGAAGAGCAAGGTGGCA 1472

DB 1148 CCCCAGCACCTGAGTTTAACTTCTCAAGAGATTGATTACTTGAAGATTAAGTGAG 1207  
QY 1473 GCAACTTTGTTAGAGAGAGAGCAATGAAGAGAGATGAGAAATTCAGGTGAGA 1532  
DB 1208 CCTCTTTTGTAAAGAGAGAGAGCTATGAAAAAGAGATTATGAAGAGTACAAAATTGCA 1267  
QY 1533 ATCAATGCTTATGTTGCCAAGCCAGAAAGTTCCCTGAAGAGATGACATGCAAGAT 1592  
DB 1268 ATCAATGCTTATGTTGCAAGGCTCAGAAAACACCTGATGAAGATGACATGCAAGAT 1327  
QY 1593 GGAACCCCTGCGCTGGAACAATGTTCTGATCATCTCGAATGATTGAGTCTTCTT 1652  
DB 1328 GGAATCTTGGCCAGAAATTAACCCGCGTATCACCTGGCATGATTGAGGTTTCTT 1387  
QY 1653 GGCCAAAGCGAGCGCTTGAATGAGGAAATGAATGAACTGCCAGATTGTTATGTTCT 1712  
DB 1388 GGATATAGTGTGCTCGTACATCGAAGAAATGAATGAACTTCTCGACTGTTTACGTTCT 1447  
QY 1713 AGAGAAAGAGACAGCAGCTATTAACATCATATAAGAAAGCTGCTATGATGATGCTC 1772  
DB 1448 AGAGAGAGAGAGACCTGCTACCAACACCAAAAAGGCTGCTGAAATGCTTTGCTT 1507  
QY 1773 CGAGTCTCTGCTGTACTATAACAATGCTCCATATTTGTTAACTTGATGATGATCTAC 1832  
DB 1508 AGGTGTCTGCAAGTTCTTAACAATGCTCCCTCATCTTGAATGATGATGATGATGAT 1567  
QY 1833 ATCAACAACAGCAAGGCTATTAAGAAAGCAATGTTTATGATGACCCCTTACTAGGA 1892  
DB 1568 GTTAACAATAGCAAGGCAAGTAAAGGAGGCAATGCTTCTTGAATGAGACCAAGTTGCT 1627  
QY 1893 AAGAGGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1952  
DB 1628 CGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687  
QY 1953 TATGCTAACCGGAATGTTGTTTGTATATCAACATGAAGTTTGGATGATGATGATGAT 2012  
DB 1688 TATGCTAACAGAAACAGATTTCTTGTATGATGATGATGATGATGATGATGATGATGAT 1747  
QY 2013 GGTCAATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2072  
DB 1748 GGGCCAGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1807  
QY 2073 GCGCCCAAAACAAAGAGCCCATCAAGACTTGAACCTGCGCCCAAGTGTGCTTT 2132  
DB 1808 CCACCTTCAATGCCAA-----GTTTCCCAAGTCAATCCTCC 1843  
QY 2133 TGTGTTGCTGCTTGGCAATGAGAACAAAGAGACTAACCAACCAAAACAGAGAA 2192  
DB 1844 TCATCTTCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1890  
QY 2193 AAAAGTTATTTTCAAGAAAGAGAAACCAATCCCTGCATATGCTTGTGTA 2252  
DB 1891 -----ATCAGAGCTTTATAGGATGCAAAACGGGAAGAACTGATGCTGCATCTTAAC 1945  
QY 2253 ATTGACGAGCTGCTCCAGAGCTGAGAAATGAAGAGCCCGTATTTGAATCAACAAAA 2312  
DB 1946 CTTAGGGAATTTGACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 2005  
QY 2313 TTGAAGAAAGAAATTTGGCAATCTTCTGTTTGTACATGCACTTCTCGAAGATGCT 2372  
DB 2006 TTTGAGAAACCTTTTGGCTTATCTTCACTTCAATGAACTTCAATGAGAAATGGA 2065  
QY 2373 GGAACCTTGAAGAGTGCAGTCTGCTTCTTTTGAAGAGCTATACATGATGATGATGAT 2432  
DB 2066 GAGTGTGTAATCTGCAACCTTCCACACTAATCAAGAGCAATTCATGATGATGATGATG 2125  
QY 2433 TGTGTTATGAAGACAGACAGTGGGAAAAAGATTGGCTGATCTATGATGATGATGAT 2492  
DB 2126 TGTGCTATGAAGAGAGACTGCTGAGGAGAAAGATTGATGATGATGATGATGATGATGAT 2185  
QY 2493 ACAGAGATATTTCTAAGTGTGTTCAAGATGATGATGATGATGATGATGATGATGATGAT 2552  
DB 2186 ACTGAGATATCTTAACCGGCTTCAAAATGCACTGCCGAGAGATGAGATGATGATGATGATG 2245

QY	2553	ATACCTAAACGGGTTGCATTCAAAGGCTCTGCACCTCTGAATCTTTCAGATCGTCTTCAC	2612
Db	2246	ATGCCCTTAAGGCCAGCATTCAAAGGATCTGCACCCATCAATCTGTGATCGGTTGCAC	2305
QY	2613	CAGGTGCTTCGGTGGGCTCTGGGCTATTTAGATCTTCTCAGCAATCATTGCCCTCTT	2672
QY	2673	TGGTATGGGTAT--GGTGGCGGCTGAAATTTTGGAAAGATTTTCTCATCTCACTCC	2729
Db	2366	TGGTATGGCTTTGGAGGTGGTGGCTTAAATGGCTTCAAGAAGCTAGCATATATAACACC	2425
QY	2730	ATCGTATACCTTGGACATCTATTTCCCTCTTGGCTTACTGTATCATTTGCCCTGCATCTGT	2789
Db	2426	ATTGTCTATCCTTTCACATCCCTTCCACTCATTTGCCCTATGTTCCTACTACCAAGCAATCTGT	2485
QY	2790	TTATTGACAGGGAATTTATCACTCCAGAGCTGAATAATGTTGCCAGCCTGTGTTCATG	2849
Db	2486	CTTCTCACAGGAAATTTATCATACCAACGCTCTCAAACTGGCAAGTGTCTCTTTCTT	2545
QY	2850	TCACTTTTATCTGCATTTTGTGCTACGAGCATCTTAGAAATGAGATGAGTGGTGTGGA	2909
Db	2546	GGCCTTTCTCTTCCATTATCTGACTGCTGTCTCGAGCTCCGATGAGTGGTGTGAGC	2605
QY	2910	ATTGATGACTGTGAGGAATGAGCAGTTCTGGGTCACTTGAAGGTGTGTCTTCAACCTC	2969
Db	2606	ATTGAGGACTTATGCGCTAACGAGCAGTTTGGGTCAATCGGTGGCTTTCAGCCCATCTC	2665
QY	2970	TTTGCTGTGTCCAGGACTCTCAAGTCATAGCTGTGTGTGATACAAAGCTTCAACGCTG	3029
Db	2666	TTTGCCGCTCTTCCAAGGTTTCTTAAGATGCTTGGCGGCAATTGACACCACTTTACTGTG	2725
QY	3030	ACATCAAAAGGTGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAAATGACTACC	3089
Db	2726	ACTGCCAAAGCAGCTGATGATGACAGATTTGTGTGAGCTCTACATTGTGAATGACTACA	2785
QY	3090	TTATTGATACCTCCTACCACTTGCTTCTATTGAACCTCATTTGCTGTGCTGCGTTC	3149
Db	2786	CTTCTAATCCCTCCAAACAACCTCCTCATGTCACAATGTTGTGTGTGCGCGAATTC	2845
QY	3150	TCAATGCGATCAATAACGATATGAGTCATGGGGCCCTCTTTGGGAAGCTAATCTTT	3209
Db	2846	TCCGATGCCCTCAACAAGGGTACGAAGCTTGGGGACCACTTTTGGCAAAAGTGTCTTT	2905
QY	3210	GCAATTTGGGTGATTGTTCATCTTTATCCCTTTCTCAAAGGTTTGGTTGGAAGCAAAAC	3269
Db	2906	TCCCTTGGGTATCTCCATCTTATCCATTCTCAAAAGTCTTATGGAGCGCAAAAC	2965
QY	3270	AGGACACCAACGATGTGATCGTCTGCTGCATTTCTGCTGCTTCAATCTTCTGCTCTT	3329
Db	2966	AGGACACCAACCAATGTGTGTCTTGTGTCAGTGTGTGTGCTTCTCTCTTGT	3025
QY	3330	TGGGTTCCGATTGATCCTTCTCTTCCGAAGGATGATGG	3367
Db	3026	TGGGTTCCGATCAACCCGTTTGTTCAGCACCGCCGATAG	3063

```

RESULT 7
US-09-221-013A-1
; Sequence 1, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Belzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24

```

```

; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2248
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1887)
US-09-221-013A-1

```

Query Match	22.5%	Score 884.6;	DB 4;	Length 2248;
Best Local Similarity	68.5%	Pred. No. 6.2e-241;		
Matches 1277; Conservative	0;	Mismatches 559;	Indels 27;	Gaps 3;

QY	1495	GAGCAATGAAGAGAGATGATGAGGAATTCAAAGGTGAGAAATCAATGCCTTAGTTGCCAAG	1554
Db	2	GAGCTATGAAGAGAGATGATGAGGAAGTTTAAAGTGAGATAAATGCTCTTGTGGCCAAAG	61
QY	1555	CCCAGAAAAGTTCCTGAAGAAGATGACAATGCAAGATGGAACCCCTGGCCTGGAACA	1614
Db	62	CACAGAAAATCCCTGAGAAAGCTGACAATGACAGATGGTACTCCCTGGCCTGTAACA	121
QY	1615	ATGTCGTGATCATCCTGGATGATTCAGGTCCTTCCTGGCCAAGCGGAGCCTTGACT	1674
Db	122	ACACTAGAGATCATCCTGGATGATACAGGTGTTCTTAGGCCATAGTGGGGCTTGATA	181
QY	1675	GTCAGGGAATGAAGTCCCAAGATTGGTTATGTTTTCTAGAGAAAAACGACCAGGCTATA	1734
Db	182	CCGATGGAATGAGCTGCTTAGACTCATCTATGTTTTCTGTGAAAAAGCGGCTTGATTTC	241
QY	1735	ACCATCATAGAAGAGCTGGTGTCTATGAATGCATTGGTCCAGTCTGTGTTACTAACAA	1794
Db	242	AACACCACAAAAAGGCTGAGCTATGAATGCATGCATCCGTGTATCTGCTGTTTACCA	301
QY	1795	ATGCTCCATATTTGTTAACTTGATTTGTGATCACTACATCAACAACAAGAGCTATTA	1854
Db	302	ATGAGCATATCTTTGAACGTGATTGTGATCATTTACTTTAATAACAGTAAGGCTATTA	361
QY	1855	AGGAAGCAATGTTTATATGATGAGCCCTTACTAGGAAGAAGTTTGCTATGTACAGT	1914
Db	362	AAGAAGCTATGTGTTTCATGATGAGACCCGGCTATTGGAAGAAGTGCTGTATGTCCAGT	421
QY	1915	TCCCTCAAGATTTGATGSGATTTGATCGCCATGACCCGATATGCTAACCCGGAATGTGTCT	1974
Db	422	TCCCTCAACGTTTGAACGTATTTGATTGCAACGATCGATATGCCAACAGGAATATAGTCT	481
QY	1975	TTTTGATATCAACATGAAGSITTTGGATGTTATTCAGGGTCCAATTTATGTTGTTACTG	2034
Db	482	TTTTGATATTAACATGAAGGGGTTGATGTTATCCAGGTCCAGTATATGTGGTACTG	541
QY	2035	GATGTGATTAGAAGGCAAGCATTTATATGTTATGATGCCCCCAAAACAAGAAGCCAC	2094
Db	542	GTTGTTGTTTAAATAGCAGGCTCTATATGGGTATGATCTGTTTTGACGGAAGAAGATT	601
QY	2095	CATCAAGGACTTGCAACTGTGCGCCCAAGTGTCCTTTGCTGTGCTGCTTGGCAATA	2154
Db	602	TAGAACCAATATTA-----TGTCAAGAGCTGTGCGGGTCAA	640
QY	2155	GGAAGCAAAAGAGACTACCAAAACCAAAACAGAGAAGAAAAGTTATTATTTTCAAGA	2214
Db	641	GGAAGAAAGTAAAGTAGCAA---GAAGTATTACTACGAAAAGAGAGAGGCATCAACA	697
QY	2215	AAGAAGAACCAATCCCTGTCATATGCTCTTGTTGTAATTGACGAAGCTGCTCCAGAG	2274
Db	698	GAAGTGACTCCAATGCTCCACTTTTCAATATGAGAGACATGATGAGGGTTTGAAGGTT	757
QY	2275	CTGAGAATGAAAAAGGCCGGTATTTGTAATCAACA AAAATTAGAAAAAGAAATTTGCCAAT	2334
Db	758	ATGATGATGAGAGGTCTATTCTAATGTCCACAGAGAGATGTAAGAAAGCGTTTGTGTAGT	817





Db 656 GTTCCTGTGAGGATGTGGAACCCCTCCAGGACTTGAAATTCCTATGGGATTAACAGTGT 715  
QY 771 TATGGTATGTAGTGTCTTGGAAGAACGGATGGAGATTTGGAAGCAGACAAAG 830  
Db 716 GACTGGCAAGAAAGTTGCCAGCTGGAGAACAGACAAAATATGATGACGTA 775  
QY 831 AGGATGCACAGACGGGGAATGATGTGTGTGATGATG-----TGACGATGCT 881  
Db 776 GCTAATAATATCCAGAGGCAAGAGGGGAGACATGGAAGGACTGTTCAATGCTGAA 835  
QY 882 GATCTACCACTAATGATGAAGCAAGACAACTGTCAGGAAATTCACCTTCATCA 941  
Db 836 GATATCCAAATGTTGATGATGACAGCTACCTGTGAGCCGACATAGTGCCATCCCTTCA 895  
QY 942 AGCCAGATTAATCCATATAGATGATATCATTAATCGGCTTGTGTTGGGTTCTTC 1001  
Db 896 AACCACTCAACCTTAACCGGATGTTATCATTCCTCGTCTTATCATCCTGATGTTCTTC 955  
QY 1002 TTCCACTACCGAGTGATGATCCGGTGAATGATGATTTGCTTGTGGCTCATCTGTT 1061  
Db 956 TTCCAATATCGTGTACTCATCCAGTGCGGAGTGTATGATGTGTGCTAGTATCTGTT 1015  
QY 1062 ATCTGTGAATCTGTTTGGCATGTCTTGATTTGATCAATCCCAAGTGTCTCCT 1121  
Db 1016 ATCTGTGAATTTGGTTGCCCTTATCCTGTGCTCTAGATCAATCCCAAGTGTACCCG 1075  
QY 1122 ATTGAGAGAGACTTAAGTCTAGACCGGCTGTCACTGAGGTTGCAAGGAGGCCAGCCA 1181  
Db 1076 ATAAACCGTGAACATACCTTGACAGGCTTGCAATGAGATATGATAGGAGGAGAGCCA 1135  
QY 1182 TCTCAACTGTCTCAATGATTTCTTGTGATGAGGTTGATCCCTAAAGAACTCTCCT 1241  
Db 1136 TCACAGCTGTCTCCATGATGTTGTGATGAGGATGATCCACTAAAGAACTCTCCT 1195  
QY 1242 TTGTCACAACAATACTGTCTATCTATCTCTTCCGCTGATATCTGTTGAATAAGTT 1301  
Db 1196 CTGATCACAGCAAACTGTTTGTCCATTTGCTGTGATTAACCTGTGACAAAGTG 1255  
QY 1302 TCTTGTATGTTCTGATGATGCTGCTCAATGCTAACTTTGAAGCATTAATCTGAAACA 1361  
Db 1256 TCATGTATGTTTGTGACGATGTTGATGTTAACTTTGAGGCTCTGTGAGAAACT 1315  
QY 1362 TCTGAATTTGCAAAAGATGGGTTCTTCTGCAACGCTACAATATTTGAACCTCGGCT 1421  
Db 1316 GCAGAAATTTGCTAGAGATGGGTTCCGTTTGGCAAGACACAATATTTGAACGAGCT 1375  
QY 1422 CCAGAGTGTACTTCCAACAGAGATAGACTACTTGAAGAACAAGTGCGCAAACTTT 1481  
Db 1376 CCAGAGTTTACTTGTCTCAAAAATAGATTACCTGAAGACAAATCCACTTCTCTT 1435  
QY 1482 GTTAGGAGAGAGAGCAATGAAGAGAGATGAGCAATTCAGGTGAGATCAATGCC 1541  
Db 1436 GTTAAAGAAAGCGGGCAATGAAGAGAGATGAAGAAATTCAGGTACCGATCAATGCT 1495  
QY 1542 TTAGTTGCCAAAGCCCAAGAAAGTTCTTGAAGAGATGGAACAATGCAAGATGGAACCCC 1601  
Db 1496 CTGTGTGCAAGGCAAAAAGTACCTGAAGAGGAGTGAACCATGCTGATGGCACTGCT 1555  
QY 1602 TGGCCTGAAAACAATGTTCTGTGATCATCTGGAATGATTCAGGCTTCTTGGCCAAAGC 1661  
Db 1556 TGGCCTGGAATTAACCAAGGATCACCTGCGATGATTCAGGTGTTCTTGGGGCACAGT 1615  
QY 1662 GGAGGCTTGACTGTGAGGAAATGAACCTGCCAGATTGTTTCTAGAGAGAAA 1721  
Db 1616 GGTGGCTTGAACATGATGTTAAGAGTGTGCAAGGCTTGTCTACGTCTCTGTAAG 1675  
QY 1722 CGACCAAGCTATTAACATTAAGAAAGCTGTGCTATGATGATGATGCTGCTGAGTCTCT 1781  
Db 1676 AGGCCAGGATTCAGATCATCAAGAAGGCTGTGCAATGAATGATGATGCTGATCT 1735  
QY 1782 GCTGT 1786  
|||||

Db 1736 GCTGT 1740  
RESULT 9  
US-09-221-013A-3  
; Sequence 3, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; PRIOR APPLICATION NUMBER: 1998-12-23  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 8411  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-221-013A-3  
Query Match 10.1%; Score 397.8; DB 4; Length 8411;  
Best Local Similarity 54.1%; Pred. No. 4.2e-102;  
Matches 1540; Conservative 0; Mismatches 657; Indels 648; Gaps 10;  
QY 1158 AGGTCACAAGGAAGGCCAGCCATCTCACTTGCTCCAAATGATTTCTTGTGATGACG 1217  
Db 4781 AGATATGATGAGACGGTGAACCATCAACAGCTCGTTCCTGTGATGTTGTTAGTACA 4840  
QY 1218 GTTATCCCTTAAAGAACCTCTTGGTCAACAACAATACTGTTCTATCTATCCTTTCG 1277  
Db 4841 GTGACCCCATTTGAAGAGCCTCCCTTGTATACAGCAACAACAGTTCTCGATTTCTTCT 4900  
QY 1278 GTGATATTCCTGTTGATTAAGTTCTTCTATGTTTCTGATGATGCTGCAATGCTA 1337  
Db 4901 GTGACATACCGGTAGATTAAGTAGCTGTATGTTTCAAGATGATGTTCAAGTATGCTT 4960  
QY 1338 ACGTTGAACATTAATCTGAACATCTGAATTTGCAAGAAATGGGTTCTTCTGCAAA 1397  
Db 4961 ACCTTGAATCCCTTCTGAAACCGCTGAGTTGCAAGAAATGGGTTCAATTTGCAAG 5020  
QY 1398 CGGTACAATTTGAACCTCGCGCTCAGAGTGTACTTCCAAGAGATAGACTACTTG 1457  
Db 5021 AAATTCACATTTGAACCTAGGGCCCTGAATTTCTATTTGGCCAGAGATAGATTACTTG 5080  
QY 1458 AAAGACAAGGTGGCAGCAAACTTGTAGGAGAGAGAGCAATGA----- 1503  
Db 5081 AAGACAAGATCCCAACCGTCTTGTAAAGAGCAGCAGCTATGAAGTCAATTGAAAA 5140  
QY 1504 ----- 1503  
Db 5141 GTCCACCTGCTTCTCATCCATACGGCAAAAGATTGACTGACTTTTCTTGGTTGTAT 5200  
QY 1504 ----AGAGAGATGAGGAATTCAGGTGAGAAATCAATGCTTATGTCCAAAGCCAG 1559  
Db 5201 TGACAGAGAGATGAGAGATTAAAGTGAAGATTAATGCTCTTGTGCCAAAGCACAG 5260  
QY 1560 AAAGTTCTGAAGAGATGAGATGCAAGATGGAACCCCTGCGCTGGAACAATGTT 1619  
Db 5261 AAAATCCCTGAAGAGCTGGAATGCAAGATGTAATGTAATGTAATGTAATGTAATGTA 5320  
QY 1620 CGTATCATCTCTGAATGA----- 1638  
Db 5321 AGAGATCATCTTGAATGATACAGGTACAGTGTGCAATCCCTTGAATTTGACAGAGAGG 5380  
QY 1639 -----TTCAAGTCTTCTTGGCCA 1658

Db 5381 ATACGTAAGAAACATGTTACATCGTTTGTTCATTTTCAGSTGTTCTTAGGCCAT 5440  
Oy 1659 AGCGAGCGCTTGACTGTGAGGGAATGAACGCCAGATTGGTTATGTTTCTAGAGAG 1718  
Db 5441 AGTGGGGGTGGAATACCAGTAGGAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGA 5500  
Oy 1719 AAACGACCAAGCTATAACCATCATAGAAGCTGTGCTATGAATGCAATGG----- 1770  
Db 5501 AAGCGGCTTGATTTCAACACACCAAAAAGCTGGAGCTATGAATGCAATGGTTGTAA 5560  
Oy 1771 ----- 1770  
Db 5561 CTTTCAGAAATCCTATTTGTGCTCTATTTTATCTCTGTTCACTGCTAAGAAAGTTTC 5620  
Oy 1771 ----- 1770  
Db 5621 TTCTGTGTAGCCGTTGCTTCACATTCCTTTTTTCTAGGCTATGTTCTCTCCTAATT 5680  
Oy 1771 -----TCCGAGTCTGTGCTGTACTACAAATGCTCCATATT 1807  
Db 5681 TAGTATCTCTTTACTTTGACAGATCCGTTATCTGCTGTTCTTACCAATGAGCATATCT 5740  
Oy 1808 GTTAACTTGATGTGTGATCATTACATCAACAAGCAAGGCTATAAGGAAGCAATGTG 1867  
Db 5741 TTGAACGTGATGTGTGATCATTACTTTAATAACAGTAAGGCTATTAAAGAGCTATGTG 5800  
Oy 1868 TTTTATGATGACCCCTTACTAGAAAGAGGTTGCTATGTAGAGTCCCTCAAGATT 1927  
Db 5801 TTTTCATGATGACCCGCTATTGAAAGAGTGCTGCTATGTCCAGTTCCTCAACGTTT 5860  
Oy 1928 TGATGGATGATCGCCATGACCGATATGCTAACCGAATGTTGTTTTTGAT----- 1982  
Db 5861 TGACGGTATGATTTTGACAGATCGATATGCCAACAAGATATAGTCTTTTCGATGTAG 5920  
Oy 1983 ----- 1982  
Db 5921 TATCACTTCCCATGTGCTTTGTTCTCTTTGTTCATATTTGGTTGATTACTCGT 5980  
Oy 1983 -----ATCAACATGAAGGTTGG 2001  
Db 5981 TTCTGCTATGCGCTGACTTGATATTTGTTCTCTTGCGCAGATTAACATGAAGGGTTGG 6040  
Oy 2002 ATGTATTCAGGGTCCAAATTATGTGTACTGATGTATTTAGAAAGCAGCATTAT 2061  
Db 6041 ATGTATCCAGGGTCCAGTATATGTGGTACTGTTGTTTAAATAGGCAGGCTCTAT 6100  
Oy 2062 ATGTTATGATGC-----CCCCAAAACAAGAAGCCACCATCAAGGACTTGCACTGCT 2115  
Db 6101 ATGGTATGATTCCTGTTTGAACGAGAAGATTAGAACCATAATATTATGTCAAGAGCT 6160  
Oy 2116 GGCCCAAGTGTGCTTTTGTGCTGTGCTTTGGCAATAGSAGCAAAAGAGACTACCA 2175  
Db 6161 GTTGGGGGTCAAGAAAGGTAAAGTAGCAAGAAGTATTACTACGAAAAGAGAGAG 6220  
Oy 2176 AACCCAAAACAGAGAAGAAAAAGTTATTTTTCAGAAAGAGAGAACCA----- 2227  
Db 6221 GCATCAACAGAGTGACTCCAATGCTCACTTTCAATATGAGGACATCGATGAGGGTT 6280  
Oy 2228 -----ATCCCTGCAATAGCT 2243  
Db 6281 TTGAAGTTTGAATTGAGCTGATTTGTATAACATCACTTCTTATGTATGATTATGT 6340  
Oy 2244 CTGTGTAATTCAGCA-----GCTGCTCCAGAGCTGAGAAAGCAAGCCGATATTG 2298  
Db 6341 GATGTTGAATCTTACAATCCTGTTTATGACAGTTATGATGATGAGAGGCTATTTCTAA 6400  
Oy 2299 TAAATCAACAAAATTAGAAAAGAAATTTGCCCAATCTTGTGTTTGTACATCCACAC 2358  
Db 6401 TGTCCAGAGAGTGTAGAGAGCGTTTGTGTCAGTGCCTGGTATTTATTGCGGCAACCT 6460  
Oy 2359 TTCTCGAAGATGTTGGAACCTTGAAGAGTCAAGTCTGCTTCTTTTGAAGAAGCTA 2418

Db 6461 TCATGGAACAAGCGGCAATCCACCAACAACCAATCCCGCTACTCTTCTGAAGAGGCTA 6520  
Oy 2419 TACATGTCATTAGTTGTGTTATGAGACAAGACAGACTGGGAAAAG----- 2466  
Db 6521 TTCATGTTATAAGCTGTGTTACGAAGACAAGACTGAATGGGGCAAGAGTCACTTTTC 6580  
Oy 2467 ----- 2466  
Db 6581 AAATGCAGCTACAGAACTTCTTATGTTCTCTTCTTACCCTGTTGATGACATCTTATTT 6640  
Oy 2467 -----AGATTGCTGATCTATGATCAGTTACAGAAAGATATTCTAATGCTGTT 2514  
Db 6641 GGCACCTTTGTTAGATTGTTGATCTATAGTTCCGTGACGGAAGATATTCTTACTGGGT 6700  
Oy 2515 TCAAGATGCATTGTCAATGTTGCGGTCATTTACTGCATACCTTAAACGGGTGCATTCA 2574  
Db 6701 TCAAGATGCATGCCGGGGTTGATATCGATCTACTGCAATCCTCCACGCCCTGCGTTCA 6760  
Oy 2575 AAGTTCTGCACCTTGAATCTTTCAGATCGTCTTCAACAGGTGCTTGGTGGCTCTTG 2634  
Db 6761 AGGATCTGCACCAATCAATCTTCTGATCGTTGAAACCAAGTCTTCGATGGGCTTTGG 6820  
Oy 2635 GGTCTATTGAGATCTTCTCAGCAATCATTTGCCCTCTTGGTATGGGTATGGTGGCGTTC 2694  
Db 6821 GATCTATCGAGATTCTTCTTAGCAGACATGTCTTATCTGTTATGTTACATGGAAGGT 6880  
Oy 2695 TGAATTTTGGAAAGATTTTCTACATCAACTTCATCGTGTATCTTGGACATCTATTTC 2754  
Db 6881 TGAGACTTTTGAGAGAGATCGCTTATATCAACACATCGTCTATCTTATTAATCAATCC 6940  
Oy 2755 CCTCTTGCTTACTGATCAATTCCTGCTGCATCTGTTTATGACAGGAAATTTATCACTC 2814  
Db 6941 CTCTTATGCGTATGTAATCTTCCCGCTTTTGTCTCATCACCGACAGATTCATCATAC 7000  
Oy 2815 CAGAGCTGAATTAATGT----- 2831  
Db 7001 CCGAGGTTTGTAAACGACCACTGCTATTATTAATTGAATCCCATTTTGTGAATGC 7060  
Oy 2832 -----GCCAGCCTGTGTTCAATGTC 2852  
Db 7061 ATTTTTTGTATCATCATGATTTGTTGACAGATAAGCACTACGGAGATTTTGGTTCACTCTA 7120  
Oy 2853 CTTTTTATCTGCATTTTGTCTACGAGCATCCCTAGAAAATGAGATGAGTGTGTTGAAAT 2912  
Db 7121 CTCTTATCTCAATTTGCTGTGACTGGAATCTTGAGCTGAGATGAGCGGTGTGAGCAAT 7180  
Oy 2913 GATGACTGTGAGGAATGAGCAGTTCTGGGTCAATGGAAGGTGTCTTACACCTCTTT 2972  
Db 7181 GAGGATTTGTGAGAGACGAGCAGTTCTGGGTCAATGTGGCACATCCGCCATCTTTT 7240  
Oy 2973 GCTGTGTTCCAGGACTTCTCAAGGTCAATAGCTGTGTGTATACAAAGCTTCAACGTGACA 3032  
Db 7241 GCTGTCTTCCAAAGGTCTAATAAGGTTCTGTGTATGACACCAACTTCAACGTTACA 7300  
Oy 3033 TCAAGG---GTGAGATGATGAGGAGTTCAGAGCTATATACATTTCAAAATGACATAAC 3089  
Db 7301 TCTAAAGCCACAGAGAGATGGGATTTTGAGAACTCTACATCTTCAAAATGACAGCT 7360  
Oy 3090 TTATTGATACCTCCTACACCTTCTCTATTTGAATTCATTGTTGTGTGCTGCGCTT 3149  
Db 7361 CTCTCATTTCCACCAACCAACCGTCTACTTGTGAACCTCATAGGCAATGTGCTGTGCTC 7420  
Oy 3150 TCAATTCGATCAATTAAGGATATGAGTCATGAGGCCCCCTCTTTGGGAAGCTATTCTTT 3209  
Db 7421 TCTTATGCTGTAACAGTGGCTACAGTCTGAGGCTGCTTTTGGGAAGCTCTTCTTC 7480  
Oy 3210 GCATTTGGGTGATGTCATCTTATCCCTTCTCAAAAGGTTGTGGAAGCAAAAC 3269  
Db 7481 GCTTATGGGTATATGCCATCTCTACCCCTTCTGAAAAGTCTGTGGAAGCAAAAC 7540  
Oy 3270 AGGACCAACGATTTGTCATGCTGTGCTCATTTCTGTGCTTCAATCTTCTGCTCCTT 3329  
Db 7541 CGAACACCAACCATGTCATTTGCTGTGCTGTTCTTCTGCGCTCCATCTTCTGTTGCTT 7600

QY 3330 TGGGTTCGATGATGATCCTTCTTG 3354  
Db 7601 TGGGTACAGATCAATCCCTTGTGG 7625

RESULT 10  
US-09-221-013A-4  
; Sequence 4, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: PCT/AU97/00402  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 5009  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(5009)  
; OTHER INFORMATION: N is A, T, G or C.  
US-09-221-013A-4

Query Match 8.0%; Score 316.8; DB 4; Length 5009;  
Best Local Similarity 53.5%; Pred. No. 3.2e-79;  
Matches 1384; Conservative 0; Mismatches 542; Indels 662; Gaps 8;

QY 894 ATGATGAAGCAAGACAACTGTCCAGAAATTCACCTTCATCAAGCCAGATTAT 953  
Db 2422 ATGATGAAGGAAGAACCTCTCTCAAGAAAGCTACTATTGTTCAAGCAGAAATTAAT 2481  
QY 954 CCATATAGATGATTATCATTAATTCGGCTTGTGTTGGGTTCTTCTTCCACTACCGA 1013  
Db 2482 CCTTACAGATGTTAATCTGTGTGCGCTCGCATTTGTTCTTTCTTCAATTATAGA 2541  
QY 1014 GTGATGCATCCGGTAATGATGATTTGCTTGTGCTCATATCTGTATCTGTGAATC 1073  
Db 2542 ATTCTCCATCCAGTCAATGATGATGATTATGTTAAAGTCAGTTATATGCGAGATA 2601  
QY 1074 TGGTTTGCATGCTTGTGATTTCTTGATCAATTCCTAAGTGTTCCTTATGAGAGAG 1133  
Db 2602 TGGTTTGCAGTGTCTTGTGATTTCTTGATCAATTCCTCAAAATGCTATAGAACGTGAA 2661  
QY 1134 ACTTACCTAGACCGGCTGTGAC-----TGAGGTTGCA 1165  
Db 2662 ACATACCTGATAGACTCTCTCTCAGGTAAACATAAACCCTGAAAAGTTCTGTCTGAAA 2721  
QY 1156 -----TGAGGTTGCA 1165  
Db 2722 TATTCATTTTTCATTTCCAAAATTTTGAACCTCTATTTTCTTACATAGGTAGCA 2781  
QY 1166 CAAGGAAGCCAGCCATCTCAACTGCTCCAAATGATTTCTTGTGAGTACGTTGATCC 1225  
Db 2782 GAAAGGAAGAAAACCGTCAGAGATTAGCACCTGTTGATTTTGTAGTACAGTGATCC 2841  
QY 1226 CTTAAAGGAACCTCTTGTGTCACACAAATATCTTATCTATCTTCCGTTGATTA 1285  
Db 2842 GTTGAAGAGCCACCTTGATTAACGAAACACAGTTCTTCCATTTAGCAGTTGATTA 2901  
QY 1286 TCCTGTGATAAGGTTCTTGTCTATGTTCTGTGATGATGCTGCAATGCTAACGTTTGA 1345

Db 2902 TCCTGTGATAAGGTTGCGTGTATATGATCAAACAATGCTGACGTATGCTTACATTTGA 2961  
QY 1346 AGCATTAATCTGAACAATCTGAATTTGCAAGAAATGGGTTCTTTCTGCAACGGTACAA 1405  
Db 2962 AGCTCTCTGATACAGCTGAGTTTGTCTAGAAAATGGGTTCTTTTGTAGAAAGTTTAA 3021  
QY 1406 TATTGAACCTGCGCTCCAGAGTGTACTTCCAAAGAGATAGACTACTTGAAGACAA 1465  
Db 3022 TATCGAGCCACGAGCTCCGAGTGTATTTTCTCAGAAAGATGATTAATCTGAAAGACAA 3081  
QY 1466 GGTGCGACCAACTTTGTTAGGAGAGAGACGAATGA----- 1503  
Db 3082 AGTTCACTCGCTTTTGTCAAGGAACGTCGTCTATGAAGGTTTCTTGTGCTTTTTC 3141  
QY 1504 ----- 1503  
Db 3142 TCTTTCTGATATATCTCTATCATAAAGTGTGTTTCAAGAAATCGATTTACGTTTGTG 3201  
QY 1504 -----AGAGAGATAGAGAAATTCAGGTGAGAAATGACCTTAGTT 1547  
Db 3202 CTGTTGTTGTTGTCAGAGAGATTAAGAGTTTAAAGTGAAGATTAATGCACTGTT 3261  
QY 1548 GCCAAGCCAGAAAGTTCTCTGAAGAAGATGACAATGCAAGATGAAACCCCTGGCCT 1607  
Db 3262 GCTACTGCACAGAAAGTCCCTGAGGAAGTTGACTATGCAAGATGGAACCTCTTGCCCT 3321  
QY 1608 GGAACAATGTTCTGATCATCTCTGGAATGAT----- 1640  
Db 3322 GGAACAACGTCGTCAGCAATCTCTGGAATGATTCAGGTAATGATGATTTGATGAATAG 3381  
QY 1641 ----- 1640  
Db 3382 GCAAAAAAAGCGGTTTGTGTCCTCTTCACTTGTTCCTCGAATCTGTTAAATTGGA 3441  
QY 1641 ----- 1640  
Db 3442 TGAGCACTTACTTCTCAATATATCTTTCAGACCGAAGCCTTTTAAAGATTTTGTAAAT 3501  
QY 1641 --CAGTCTTCTTGGCCAAAGCGGAGGCTTGACTGTGAGGAAATGAATGCCACGAT 1698  
Db 3502 GACAGGTGTTCTTGGGTACATAGTGAAGTGTGATACGATGTAATGATTAACCAAGTC 3561  
QY 1699 TGGTTATGTTTCTAGAGAGAAACGACGCTATTAACCATCATTAAGAAAGCTGTGCTA 1758  
Db 3562 TAGTATGTTTCTCTGAGAAAGCGGCTGGAATTTGATCACCACAAAGAACTGGAAGCTA 3621  
QY 1759 TGAATGCATTGG----- 1770  
Db 3622 TGAATTCCTTGAAGTATATGTTGTTCTTATTAATCTCTCTTTCCGAGCCCT 3681  
QY 1771 -----TCCAGTCTGCTGTAC 1788  
Db 3682 GACTTCTATAAATAAACTCATCTTACTTCTTCTGAAGATCCGAGTCTGCTGTTTC 3741  
QY 1789 TAACAATGCTCCATATTGTTAAACTTGATGTGATCATCATCAACAACGCAAGG 1848  
Db 3742 TATCAAAAGCTCTTACCTTCTTAATGTCGATTTGTGATCACTACATCAACAACGCAAG 3801  
QY 1849 CTATAAGGAAGCAATGTGTTTATGATGACCCCTTACTAGGAAAGAGGTTTGTATG 1908  
Db 3802 CAATTAGAGAACTATGTGTTTCTATGATGACCCGCAATCGGAAAGAAAGTTTGTATG 3861  
QY 1909 TACAGTTCCCTGAAGATTTGATGGATGATGCCATGACCGATATGCTAACCGGAATG 1968  
Db 3862 TTCAGTTCCGAGAGATTTGATGGATGATGACATGATAGTACTCAAAACCGTAACG 3921  
QY 1969 TTGCTTTTGTGAT----- 1982  
Db 3922 TTGTTTCTTTGATGATGTGTCCTTATCTTTTGTGTTTCTGTTTATGTTTATG 3981  
QY 1983 -----ATCAACATGAAGGT 1997  
Db 3982 GCTTTCCTCTTCTCATTTGATATGTTTGTGTTGGAACAAGTTAATGAAGGT 4041



QY	1998	TTGATGCTATTACAGGCTCAATTTATGTTGTTACTGATGTGTATTTAGAAAGCAGCA	2057
Db	4042	CTTGATGGGATACAAAGACCGATATATGTGCGGACAGTTGTGTTAGAAAAACAGCT	4101
QY	2058	TTATATGGTTATGATGCCCCCAAAACAAGAACCCATCAAGGACTTGCAACTGCTGG	2117
Db	4102	CTTTATGTTTATGATGCACCAAGAAGAAGAAACCAACGCAAACTGTAACGTGTCG	4161
QY	2118	CCCAAGTGTGCTTTTGTGTTGCTGCTTTGGCAATAGGAAGCAAAAGAAGACTACCAA	2177
Db	4162	CCTAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	4221
QY	2178	CCCAAAACAGAGAAGAAAAGTTATATTTTC	2210
Db	4222	AAGAAACTAACAATAAGAGACTTCAAGCAGATTCTATGCCCTAGAGAATGTGACGA	4281
QY	2211	-----AAGAAAGAGAACCAATCCCCCTGCATATGCTC	2244
Db	4282	GGTGTATCGTCCCAAGTTAAAAAAGAAAGGAAAAAACAATTTCTTATTTGGTTCTG	4341
QY	2245	TTGGTGAATTGACGAAGCTGCTCC-----AGGAGCTGAGAATGAAAGGCGCG	2293
Db	4342	TCTTGTGAAAGCTTAAGTAGATCTTTTGTATTGTATGTTCAATGTTGAGAAGATC	4401
QY	2294	TATGTAAATCAACAAAAATAGAAAAAGAAATTTGGCCATCTTCTGTTTGTATCATC	2353
Db	4402	TGAAGCAACACAATTGAAATTTGAGAAGAGTTTGACATCTCCGGTTTCGTTGCTC	4461
QY	2354	CACACTTCTCGAAGATGTGGAACCTTGAAGAGTGCAAGTCTGCTCTCTTTTGAAGA	2413
Db	4462	TGCTGTCTACAGAACGGTGAAGTTCCCGTAACGCAAGCCCGCATGTTGTTAAGAG	4521
QY	2414	AGCTATACATGTCATTAGTTGTTGTTATGAGAACAAGACAGACTGGGAAAG-----	2466
Db	4522	AGCCATTCAAGTTATTAGCTGCGGGTACGAAGATTAACGAATGGGAAAAAGAGTAGA	4581
QY	2467	-----	2466
Db	4582	AAACATTACAAGTTTTCAACTTCTGAAAACTAGAAAAGTTCTTGATCTCATTCCTG	4641
QY	2467	-----AGATTGGCTGATCTATGATCAGTTACAGAAATATCTTAACGTG	2511
Db	4642	CTGATTAATCACACGACAGATCGGGTGAATTTATGATCGGTGAAGATATCTGACGG	4701
QY	2512	GTTTCAAGATGATGTCATGTTGTTGCGGTCAATTATTCATGATACCTTAAACGGTTCAT	2571
Db	4702	GTTTCAAGATGATGTCATGTTGTTGCGGTCAATTATTCATGATACCTTAAACGGTTCAT	4761
QY	2572	TCAAAGGTTCTGCACCTCTGAATCTTTCAGATCGTCTTCAACAGGTGCTTCGGTGGCTC	2631
Db	4762	TTAAAGGATCTGCTCTCTTAATTAACCTGTCAGATCGTCTTCAATCAAGTTCTACGTTGGCTC	4821
QY	2632	TTGGCTTATGAGATCTTCTTCAAGCAATCATGTCCTCTTGGTATGATGATGATGATG	2691
Db	4822	TTGGCTTATGAGATCTTCTTCAAGCAATCATGTCCTCTTGGTATGATGATGATGATG	4881
QY	2692	GTTCTGAATTTTGGAAAGATTTTCTTACATCAACTCACTCGTGTATCTTGACATCTA	2751
Db	4882	GTTTAAATGTTGGAGAGATTTCTTACATCAACTCTGCTATCTTGACATCTCAC	4941
QY	2752	TTCCCTCTTGCTTACTGATCATGTCCTGCCATCTGTTTATGACAGGAAATTTATCA	2811
Db	4942	TTCCATGATCGTCTATGTTCTCTCCCGCGGTTGTTTACTCACAGGAAATTCATCG	5001
QY	2812	CTCCAGAG 2819	
Db	5002	TCCCTGAG 5009	

```

: Patent No. 6476212
:
: GENERAL INFORMATION:
: APPLICANT: Lalgudi, Raghunath V.
: APPLICANT: Ito, Laura Y.
: APPLICANT: Sherman, Bradley K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
: FILE REFERENCE: PL-0017 US
: CURRENT APPLICATION NUMBER: US/09/313,294A
: CURRENT FILING DATE: 1999-05-14
: NUMBER OF SEQ ID NOS: 7600
: SOFTWARE: PERL Program
: SEQ ID NO 849
: LENGTH: 268
:
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. 6476212 700549940H1
US-09-313-294A-849

```

```

Query Match      6.8%; Score 268; DB 4; length 268;
Best Local Similarity 100.0%; Pred. NO. 3.1e-66;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	2849	GTCACTTTTATCTGCATTTTGTGCTACGAGCATCTAGAAATGAGATGAGTGGTGTGG	2908
Db	1	GTCACTTTTATCTGCATTTTGTGCTACGAGCATCTAGAAATGAGATGAGTGGTGTGG	60
QY	2309	AATGATGACGTGTGAGGAATGAGCAGTTCTGGTCATTGGAGGTGTCTCAACCT	2968
Db	61	AATGATGACGTGTGAGGAATGAGCAGTTCTGGTCATTGGAGGTGTCTCAACCT	120
QY	2969	CTTTGCTGTGTTCCAGGACTTCTCAAGTCAATAGCTGTGTGATACAAAGTTCACCGT	3028
Db	121	CTTTGCTGTGTTCCAGGACTTCTCAAGTCAATAGCTGTGTGATACAAAGTTCACCGT	180
QY	3029	GACATCAAAGGCTGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGAGCTAC	3088
Db	181	GACATCAAAGGCTGAGATGATGAGGAGTTCTCAGAGCTATATACATTCAAATGAGCTAC	240
QY	3089	CTTATTGATACCTCTCAACCACTTGCTT	3116
Db	241	CTTATTGATACCTCTCAACCACTTGCTT	268

```

RESULT 12
US-09-313-294A-7123
; Sequence 7123, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 7123
; LENGTH: 317
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381163H1
; NAME/KEY: unsure
; LOCATION: 10
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7123

```

RESULT 11  
US-09-313-294A-849  
; Sequence 849, Application US/09313294A

Query Match	5.4%;	Score 213;	DB 4;	Length 317;
Best Local Similarity	83.3%;	Pred. No. 1.5e-50;		

Matches 265; Conservative 0; Mismatches 51; Indels 2; Gaps 2;

```
QY 2172 ACCAACCACAAACAGAGAGAAAAAGTTATTTTCAAGAAAGAGAACCAATCC 2231
    |||||
Db 1 ACAAACCAAGACGAGAGAGAAAAAGTTATTTTCAAGAAAGAGAAAAACCATCT 60
QY 2232 -CCTGCATATGCTCTTGGTGAATTCGAGAGCTCTCCAGAGCTGAGATGAAAGGC 2290
    |||||
Db 61 ACCTGCATATGCTTGGTGAATTCGAGAGCTCTCCAGAGCTGATATCGAAGAGC 120
QY 2291 CGGTATGTAATCAACAAATTAGAAGAAATTTGGCCAATCTCTGTTTTTTTGTAC 2350
    |||||
Db 121 CGGAATCGTAATCAACAGAACTAGAGAAATTTGGCAGCTCTCTGTTTTGTGCG 180
QY 2351 ATCCACACTTCTCGAAGATGTGGAACCTTGAAGAGTGCAGAGCTCTGCTTTTGA 2410
    |||||
Db 181 ATCAACACTTCTTGAAGACGAGGAGACCTTGAAGAGCGCAAGTCCAGCTTCTTGAA 240
QY 2411 AGAAGCTATACATGTCATTAGTTGTTAGTGAAGACAAAGACAGACTGGGAAAAAGAGAT 2470
    |||||
Db 241 GGAAGCTATACATGTTATCAGCTGCGGCTACGAAAGACAGACCGACTGGGAGTAAGAGA- 299
QY 2471 TGGCTGATCTATGATC 2488
    |||||
Db 300 TGGCTGATTTACGATC 317
```

## RESULT 13

```
US-09-313-294A-4015
; Sequence 4015, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4015
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700347541H1
; NAME/KEY: unsure
; LOCATION: 32, 38, 125, 162, 236, 246, 248
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4015
```

Query Match 4.2%; Score 164.2; DB 4; Length 278;  
Best Local Similarity 87.4%; Pred. No. 1.1e-36;  
Matches 201; Conservative 0; Mismatches 26; Indels 3; Gaps 2;

```
QY 2363 CGAGATGTGGAACCTTGAAGAGTGAAGT-CCTGCTTCTCTTTGAAGAGCTATAC 2421
    |||||
Db 1 CGAAATGTGGAACCTTGAAGAGTGAAGTCCGCTCTCTTGAAGAGAGCTATAC 60
QY 2422 ATGTCATAGTTGTTATGAAGACAGAGACTGGGAAAAAGATTGGCTGATCT 2481
    |||||
Db 61 ATGTCATAGTTGTTATGAAGACAGAGACTGGGAAAAAGATTGGCTGATCT 120
QY 2482 ATGATCAGTTACAGAGATATTCTAAGTTCAGATGATGTCATGTTGGC--G 2539
    |||||
Db 121 ATGATCAGTTACAGAGATATTCTAAGTTCAGATGATGTCATGTTGGC--G 180
QY 2540 GTCAATTTACTGCATACCTTAACGGGTTGCATTCAAGGTTCTGCACCTC 2589
    |||||
Db 181 TCAATTTTACTGCATACCTTAACGGGTTGCATTCAAGGTTCTGCACCTC 230
```

## RESULT 14

```
US-09-313-294A-6485
; Sequence 6485, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6485
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351882H1
US-09-313-294A-6485
```

Query Match 4.0%; Score 159; DB 4; Length 281;  
Best Local Similarity 73.1%; Pred. No. 3.2e-35;  
Matches 204; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```
QY 2720 CATCACTCCATCGTGTATCCCTTGACATCTATCCCTCTTGCTTACTGTACATGGCC 2779
    |||||
Db 1 CATCAACACCAACCATCTACCCCGCTCACGTCCTCCCGCTCTCATTTACTGTATCTGCC 60
QY 2780 TGGCATCTGTTTATGACAGGGAATTTATCACTCCAGAGCTGAATATGTTGCCAGCCT 2839
    |||||
Db 61 TGGCATCTGCTGCTCACGGGGAATTCATCATCCAGAGATCAACAATTCGCTAGTAT 120
QY 2840 GTGTTTCATGTCACCTTTTATCTGCATTTTGTCTACAGAGCATCCTAGAAATGAGATGAG 2899
    |||||
Db 121 CTGTTTCATCTCTCTCTTCATCTGCATCTTCCGACCGGTTCTTGAGATGAGGTGAG 180
QY 2900 TGGTGTGAATTGATGACTGTGTGAGGAATGAGCAGTTCTGGTTCATTGAGGTGTGC 2959
    |||||
Db 181 CGGCGTGGCATCGACGAGTGTGAGGAAGAGCAGTCTTGAGGTTCATCGAGGATCTC 240
QY 2960 CTCACACTCTTTGCTGTGTTCAGGAGCTTCTCAAGT 2998
    |||||
Db 241 CGCCCACTCTTCGCCGCTCTTCCAGGGCTCTCAAGT 279
```

## RESULT 15

```
US-09-313-294A-4753
; Sequence 4753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4753
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348982H1
; NAME/KEY: unsure
; LOCATION: 64, 189
; OTHER INFORMATION: a, t, c, g, or other
```

US-09-313-294A-4753

Query Match 3.7%; Score 145; DB 4; Length 300;  
Best Local Similarity 70.0%; Pred. No. 3.2e-31;  
Matches 208; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

```
QY      2818 AGCTGAATATGTTGCCAGCTGTGTTTCATGTCACCTTTTATCTGCATTTTGTCTACGA 2877
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       3  AGATCAGCACTTCGCCAGCACTGTTCATCTCCCTTCATCTCGATCTTCGCCACGG 62

QY      2878 GCATCCTAGAAATGAGATGAGTGTGTTGAATTGATGACTGTGAGGAATGACAGT 2937
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       63  GNATCCTGAGATGAGTGGAGCGGGGTGGCATCGACGAGTGGTGAGGAACGACAGT 122

QY      2938 TCTGGGTTCATGAGGTGTGCTCCACACCTCTTGTGCTGTTCAGGACTTCTCAAGG 2997
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      123  TCTGGGTGATCGGGGCATCTCCGCCACCTCTTCGCCGTGTTCCAGGCGCTGCTCAAGG 182

QY      2998 TCATAGCTGTGTGATACCAAGCTTCAACCGTGACATCAAAGGTTGAGATGATGAGAGT 3057
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      183  TGCTGNCGGCATCGACACCACTTCAACCGTCAACCTCCAAAGGCTCGGACGAGGACGGCG 242

QY      3058 TCTCAGAG--CTATATACATTCAAATGGACTACCTTATTTGATATCCTACCACCTT 3112
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      243  ACTTCGCGGAGCTGTACATGTTCAAGTGAGACGACGCTCCTTATCCCGCCACCAACCAT 299
```

Search completed: December 16, 2003, 00:53:10  
Job time : 248 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:00:59 ; Search time 27 Seconds  
(without alignments)  
3868.118 Million cell updates/sec

Title: US-09-720-383C-10  
Perfect score: 5918  
Sequence: 1 MEASAGLVAGSHNRNLELVI.....DFPLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4278.5	72.3	1084	2	T08583	cellulose synthase
2	4209.5	71.1	1088	2	H84604	probable cellulose
3	4202	71.0	1081	2	T52028	cellulose synthase
4	3967.5	67.0	1026	2	T51579	cellulose synthase
5	3945.5	66.7	1081	2	T05351	cellulose synthase
6	3875.5	65.5	1065	2	T52054	cellulose synthase
7	3782	63.9	1065	2	F84649	probable cellulose
8	3469.5	58.6	974	2	T10797	cellulose synthase
9	3329.5	56.3	958	2	T04870	cellulose synthase
10	2907	49.1	685	2	T10800	cellulose synthase
11	2330.5	39.4	1181	2	D86157	hypothetical prote
12	2311	39.1	1145	2	T51546	cellulose synthase
13	2256	38.1	1111	2	T05646	hypothetical prote
14	2166.5	36.6	1036	2	D84741	probable cellulose
15	2118.5	35.8	583	2	T02209	cellulose synthase
16	2114	35.7	979	2	C86446	probable cellulose
17	1056.5	17.9	757	2	T02561	probable cellulose
18	1046	17.7	712	2	T02552	cellulose synthase
19	1040	17.6	755	2	T02553	cellulose synthase
20	993.5	16.8	748	2	T02560	cellulose synthase
21	959.5	16.2	689	2	T08918	hypothetical prote
22	952	16.1	727	2	T08920	hypothetical prote
23	944.5	16.0	686	2	T08919	hypothetical prote
24	901	15.2	828	2	E71417	hypothetical prote
25	829	14.0	710	2	B71417	hypothetical prote
26	414.5	7.0	326	2	T08591	TGACG-motif bindin
27	402.5	6.8	693	2	AF2275	cellulose synthase
28	393	6.6	322	2	T12093	TGACG-motif bindin
29	359.5	6.1	322	2	T08592	TGACG-motif bindin

30	256.5	4.3	759	2	D70422	cellulose synthase
31	252	4.3	754	2	A43735	bcsA protein - Ace
32	248	4.2	874	2	AB0985	probable polysacch
33	239.5	4.0	1596	2	T31338	cellulose synthase
34	239	4.0	888	2	E91180	probable cellulose
35	238	4.0	692	2	S47754	hypothetical prote
36	238	4.0	888	2	H65151	hypothetical 101.6
37	232	3.9	888	2	G86026	probable cellulose
38	227.5	3.8	729	2	A98320	cellulose synthase
39	227.5	3.8	729	2	AD2963	cellulose synthase
40	227.5	3.8	861	2	I39714	cellulose synthase
41	211	3.6	664	2	A95889	probable cellulose
42	201	3.4	723	2	C36963	cellulose synthase
43	194	3.3	768	2	B97083	glycosyltransferas
44	175	3.0	654	2	AB2932	beta 1,3 glucan sy
45	175	3.0	654	2	D98350	hypothetical prote

ALIGNMENTS

RESULT 1						
T08583						
cellulose synthase (EC 2.4.1.-) catalytic chain - Arabidopsis thaliana						
N/Alternate names: protein T22F8.250						
C/Species: Arabidopsis thaliana (mouse-ear cress)						
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999						
C/Accession: T08583; T09014						
R/Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; M.						
submitted to the Protein Sequence Database, May 1999						
A/Reference number: Z16442						
A/Accession: T08583						
A/Molecule type: DNA						
A/Residues: 1-1084 <BEV>						
A/Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.250						
A/Experimental source: cultivar Columbia; BAC clone T22F8						
R/Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.;						
Science 279, 717-720, 1998						
A/Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.						
A/Reference number: Z13745; MUID:98111412; PMID:9445479						
A/Accession: T09014						
A/Status: translated from GB/EMBL/DBJ						
A/Molecule type: mRNA						
A/Residues: 1-1084 <ARI>						
A/Cross-references: EMBL:AF027173; NID:g2827140; PIDN:AAC39335.1; PID:g2827141						
A/Experimental source: cultivar Columbia						
C/Genetics:						
A/Gene: ATSP:T22F8.250; Ath-A						
A/Map position: 4						
A/Introns: 27/3; 156/3; 191/1; 254/2; 343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3;						
C/Keywords: cell wall synthesis; glycosyltransferase; hexosyltransferase						
Query Match						
Best Local Similarity 72.3%; Score 4278.5; DB 2; Length 1084;						
Matches 782; Conservative 127; Mismatches 161; Indels 27; Gaps 12;						
QY	1	MEASAGLVAGSHNRNLELVIRRDGPGFPPREQNGOVCOICGDDVGLAPGDPFVACNE	60			
DB	1	MNTGRLIAGSHNRNLELVIRRDGPGFPPREQNGOVCOICGDDVGLAPGDPFVACNE	60			
QY	61	CAFPVCRDCYEYERREGTQNCPOCKTRYKLGCCQRTVG-DEEDGVDDLDFNWDGHD	119			
DB	61	CAFPVCRPCYEYERREGNQACPOCKTRYKIKGSPRVDDDEEBEDIDLEVEFD-HGMD	119			
QY	120	SQSVAESMLYGHNSYGRGDPNGAPQAFQNLNPNVPLLTNGQMVDDIPBQHALVPSFMG	179			
DB	120	PEHAABAALSSRLNTGRGLDSAPP-----GSQIPLLTYCDEADMTSDRHALLVPPSTG	174			
QY	180	GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKORO-ERM---HQT	234			
DB	175	YGNRVYPAPFTSSAPPQARSMVPQKDLAEYGYGSVAMKDRMEVWTKRQGEKLVTKHEG	234			
QY	235	GNDGGG-----DDGDDADLPMLDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFPHYR	290			

Db	235	GNNGRGSNDDELDDPDMPMMEDEGRQPLSRKLP	PIRSSRINPYRMLILCRLA	ILGLFFHYR	294
Qy	291	VMHPVNDAFALWLISVCEIWFAMSWILDQFPKWP	PIERETYLDRLSRFDKEGQPSQLA		350
Db	295	ILHPVNDAYGLMLTSVCEIWFAMSWILDQFPKWP	PIERETYLDRLSRFDKEGQPSGLA		354
Qy	351	PIDFVSTVDPLKEPPLVTNTVLSISVDYPVDKVS	CYVSDGGAAMLTFEALSETSEFA		410
Db	355	PVDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVA	CYVSDGGAAMLTFEALSDTAEPFA		414
Qy	411	KKWVPFCKRYNIEPRAPEMYFOOKIDYLDKVA	ANFVRERERAMKREYEEFKVRINALVAK		470
Db	415	RKWVPFCKKFNIEPRAPEMYFSQKMDYLKKNKH	PAFVRERRAMKRDYEEFKVINALVAT		474
Qy	471	AQKVPEEGWTMQDGTWMPGNNVRDHPGMIQVFL	GSGGLDCEGNELPRLVYVSRBKRP	GY	530
Db	475	AQKVPEEGWTMQDGTWMPGNNVRDHPGMIQVFL	GSHGVRDITDGNELPRLVYVSRBKRP	PGF	534
Qy	531	NHHKAGAMNALVRVSAVLTNAPYLLNDCDHY	INNSKA	IKAMCFMMDPLLGGKVCYVQ	590
Db	535	DHHKAGAMNSLIRVSAVLSNAPYLLNDCDHY	INNSKA	IRESMCFMMDPQSGKVCYVQ	594
Qy	591	FPQRFDDGIDRHDYANRNVFFDINMKGLDGIQ	GP	IYVGTGCVFRQALYGYDAPKTKP	650
Db	595	FPQRFDDGIDRHDYANRNVFFDINMKGLDGIQ	GP	IYVGTGCVFRQALYGYDAPKTKP	654
Qy	651	PSRTCNCWPKWCFCCCCFGNRKQKTKTKTEKK	LLFFKKEENQSPAYALGEIDEA--A		708
Db	655	PGKTCNCWPKWC--CLCC--GLRKKSKT--	--KAKDKT--	--NTKETSKQIHALENVDEGVIV	706
Qy	709	PGAENKAGIVNQQKLEKFGQSSVFVTSTLLENG	T	LKSPASPLKEA	IHVISCYED 768
Db	707	PVSNVEKRSEATQLEKFGQSPVFVASAVLQNG	GVPRNASPACLLREAIQVISCYED		766
Qy	769	KTDWKEIGWIYGSVTEDILTGFKMHCHGWSIY	CI	PKRYAFKGSAPLNLSDRLHQVLRW	828
Db	767	KTEWKEIGWIYGSVTEDILTGFKMHCHGWSIY	CI	PKRYAFKGSAPLNLSDRLHQVLRW	826
Qy	829	ALGSIIEFFSNHCPWYGGGLKFLERFSYINSIV	PWTSI	P	LAYCTLPALCLTGKF 888
Db	827	ALGSVEIFLSRHCPWYGGGLKFLERFSYINSIV	PWTSI	P	LIVYCSLPVCLTGKF 886
Qy	889	ITPELNNVASLWMSLFICIFATSI	LEMWSGVIGIDMWRNEQFWVIGVSSHFAVFQ		948
Db	887	IVPEISNAGILFMLMFISIAVTGILEMQG	VGIDMWRNEQFWVIGVSSHFAVFQ		946
Qy	949	LKVIAGVDTSFVTSKGGDEEFSELYTFKWTLL	IP	PTLLLNFI	GVAGVSNAINN 1008
Db	947	LKVIAGVNTNFVTSKAADGAFSELYIFKWTLL	IP	PTLLIINI	IGVIVGVSDAISN 1006
Qy	1009	GYESWGPFLGKLFFAFVIVHLYPFLKGL	VGRNRTPTIVIVMSILLASIFSLWVRIDP		1068
Db	1007	GYDSWGPFLGRLFFALWIVHLYPFLKGL	M	GKQDKMPTIIVVMSILLASILTLLWVRNP	1066
Qy	1069	FLAKDGP	LLIEECGLDC	1085	
Db	1067	FVAK--GGP	VLEICGLNC	1082	

RESULT 2  
H84604  
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: H84604  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: H84604

A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-1088 <STO>									
A;Cross-references: GB:AB002093; NID:g4417271; PIDN:AAD20396.1; GSPDB:GN00139									
C;Genetics:									
A;Gene: At2g21770									
A;Map position: 2									
Query Match									
71.1%; Score 4209.5; DB 2; Length 1088;									
Best Local Similarity 70.2%; Pred. No. 1.7e-314;									
Matches 769; Conservative 127; Mismatches 179; Indels 21; Gaps 10;									
Qy	1	MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCQICGDDVGLAPGDPVACNE	60						
Db	1	MNTGRLIAGSHNRNEFVLINADDTARIRSAEELSGQTCIKRDEIELTDNGEPYACNE	60						
Qy	61	CAFPVCRDCEYERREGTONCPQCKTRYKRLKGCORVTDGDEEBDGVDDLNEFWMDGHS	120						
Db	61	CAFPTRPCYERYERREGNQACPQCGTRYKRIKSPVEGDEEDDIDLEHEFY--GMDP	118						
Qy	121	QSAVESML-YGHMSYGRGGDPNGAPQAFQNLNPNVPLTNGQVDDIPPRQHALVPSFMGG	179						
Db	119	EHTTEALALYMRNLNTGRGTDEVSHLYSASPGSEVPLTTYCDESDSMYSDRHALIVPSTG	178						
Qy	180	GGRIRHPLPYADPSLPVQPRSMGPSKDIAAYGYGSVAWKEEMENWKORO-ERM---HQT	234						
Db	179	LGNRVHVHPFTDSFASIHTRPMVPQKDLTVYGYGSVAWKDRMEVWKQOIEKLQVYKNER	238						
Qy	235	GNDGGG-----DDGDDADLPLMDEARQOLSRKILPSSQINPYRMIITIRLVVIGFFHY	289						
Db	239	VNDGDGDFIVDELDDPGLPMDEGRQPLSRKLPISRSRINPYRMLIFCRLAIGLFFHY	298						
Qy	290	RVNHVNDAFALMLISVCEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQL	349						
Db	299	RIHPVNDAFGLMLTSVCEIWFASWILDDQFPKYPPIERETYLDRLSLRYEKEGKSEL	358						
Qy	350	APIDFVSTVDPLKEPPLVTNTVLSISVDYPVDKVSQVSDGDAAMLTFEALSETSEF	409						
Db	359	APVDVFASTVDPLKEPPLITANTVLSILAVDYVEKVAQVSDGDAAMLTFEALSTYAEF	418						
Qy	410	AKKWPFPCKRYNIEPRAPEWYFOOKIDYLDKVAANFVRERERAMKREYEEFKVRINALVA	469						
Db	419	ARKWVPFCKKFSIEPRAPEWYFSQKMDYLKHQVDPAFVMEERRAMKRDYEEFKVINALVS	478						
Qy	470	KAQKVPPEEGWTMQDGTWPMPGNNVDRDHPGMIQVFLGSGGLDCEGNELPRLVYVSREKRP	529						
Db	479	VSQKVPEEGWTMQDGTWPMPGNNVDRDHPGMIQVFLGSHGVCMDGNELPRLVYVSREKRP	538						
Qy	530	YNHHKKAGAMNALVRSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLLGGKVCYV	589						
Db	539	FDHHKKAGAMNSLIRSAVLSNAPYLLNDCDHYINNSKAIREAMCFMMDPQSGKVCYV	598						
Qy	590	QFPQRFDDGIDRHDYANRNVFFPDINMKGLDGIQGPITYVGTGCVFRQALYGYDAPKTKK	649						
Db	599	QFPQRFDDGIDRHDYANRNVFFPDINMKGLDGIQGPITYVGTGCVFRQALYGYDAPKTKK	658						
Qy	650	PPSRTCNCWPKWCFCCCCFGNRKQKTKTKPTEKKLLFFKKEENQSPAYALGEIDEAAP	709						
Db	659	PPGRTCNCPKWC-CLCC-GMRK-KTGKVDNQK---KPKETSKQIHLEHIEBGLQ	711						
Qy	710	GAENKAGIVNQOKLEKFGQSSVFVTSTLLENGTLLKSPASLLKEAIIHVISCGYEDK	769						
Db	712	VTNAENNSETAQLEKFKFGQSPVLVASTLLNGVPSNVNPAISLLRESIQVISCYEEK	771						
Qy	770	TDWGKEIGWIYGSVTEDILTGFKMHCHGWSIYCI PKRYAFKGSAPLNLSDRLHQVLRWA	829						
Db	772	TEWGKEIGWIYGSVTEDILTGFKMHCHGWSIYCI PKRYAFKGSAPLNLSDRLHQVLRWA	831						
Qy	830	LGSIIEFFSNHCPMLWYGGGLKFLERFSYINSIVYPWTSIPLLAYCTLPALCLTGKFI	889						
Db	832	LGSVEIFLSRHCPWYGGGGLKWLERFSYINSIVYPWTSIPLLAYCSLPALCLTGKFI	891						
Qy	890	TPPELNNVASLWMSLFICIFATSILEMWMSGVIGIDMWRNEQFWVIGVSSHFAVFQGL	949						

Db 892 VPEISNAGILFLIMFSIAVTGILEMGKIGIDDWMRNEQFWIGVSSHLFALFQGL 951

Qy 950 LKVIAGVDTSTVTSTKSGDDEEFSELYTFKWTLLIPTTLILLNPIGVAGVSNAINNG 1009

Db 952 LKVIAGVSNFTVTSKADDDGEFSELYTFKWTSLIPTTLITNIVGIVGVSDAINNG 1011

Qy 1010 YESWGPFLGKLFFAFWIVHLVPLFKGLVGRQNRTPPTIVIVWSILASIFSLMVRIDPF 1069

Db 1012 YDSWGPFLGRLFFALMVIHLVPLFKGLGKQDRVPTIILVWSILASITLLMVRVNP 1071

Qy 1070 LAKDDGPLLEBEGGLDC 1085

Db 1072 VSK-DGPVLEICGLDC 1086

RESULT 3

T52028

cellulose synthase [imported] - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear creas)

C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000

C:Accession: T52028

R:Joshi, C.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z25890

A:Accession: T52028

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1081 <JOS>

A:Cross-references: EMBL:AF062485; PIDN:AAC29067.1

Query Match 71.0%; Score 4202; DB 2; Length 1081;

Best Local Similarity 70.8%; Pred. No. 6.5e-314;

Matches 770; Conservative 126; Mismatches 172; Indels 20; Gaps 10;

Qy 7 LVAGSHNRNELVIRDDPGPKPREQNGVCOICGDDVGLAPGDPFVACNECAFVPC 66

Db 4 LIAGSHNRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPPFVACNECAFVPC 63

Qy 67 RDCYEYERREGTQNCPOCKTRYKRLKGCORVTGDEEEDGVDDLNEFNW--DGHDSQSA 124

Db 64 RPCYEYERREGNOACPOCKTRFKRLKGSPRVEGEDEEDIDDLNEFEYGNNGIGFDQVS 123

Qy 125 ESMLYGHMSYG-RGDPNGAPQAFQOLNPNVPLLTNGQMVDDIPPEQHAL-VPSFMGGGK 182

Db 124 EGMGISRRNSGFPQSDLDASPPGQ---IPLTTYGDEDEVEISSDRHALIVPPSLGHHN 179

Qy 183 RIHPLPYADPSLPVQPRSMDSKDIAAYGYGSVAMKERMENWKORQ-ERMHOTGNDGGD 241

Db 180 RVHPVSLDPTVAAHRLMVPQKDLAVGYGSVAMKDRMEEMKRNQEKLOVRHHEGDPD 239

Qy 242 --DGDADLPLMDEARQQLSRKIPLESQINPYRMIIIRLVLTGFFHYRVMPVNDAF 299

Db 240 FEDGDDADFPMDDEGRQPLSMKIPKSSKINPYRMILVLRVLIGLFFHYRILHPVKDAY 299

Qy 300 ALMLISVICEIMFAMSWILDQFPKPFIERETYLDRLSLRFDKGQPSQLAPIDFVSVTV 359

Db 300 ALMLISVICEIMFAVSWVLDOFPKMYPIERETYLDRLSLRYEKKGKPSGLSPVDVFSTV 359

Qy 360 DPLKEPPLVTNTVLSILSVDPVDKVSQVSDDGAAMLTFEALSETSEFAKMWPFCKR 419

Db 360 DPLKEPPLITANTVLSILAVDYPVDKVAQVSDDGAAMLTFEALSETAEFARKWVPFCCK 419

Qy 420 YNIEPRAPEMYFOQKIDYLDKDKVAANFVERRAMKREYEEFKVRINALVAKAQKPEBGN 479

Db 420 YCIEPRAPEMYFCHKMDYLKKNKHPAFVRRERAMKRDYEEFKVINALLVATAQKVPEDGN 479

Qy 480 TMQDGTMPDGNVNRDHPGMIQVFLGQSGLDCEGNELPRLVYVGREKRPVNHKKAGAM 539

Db 480 TMQDGTMPFGNSVRDHPGMIQVFLGSDGVADVENNELPRLVYVGREKRPFGDHHKAGAM 539

Qy 540 NALVRSYAVLTVAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFQDGD 599

Db 540 NSLIRVSVGLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRFQDGD 599

Qy 600 RHDRYANRNVPFEDINMKGLDGIQGPITYGTGVFRROALYGVDAKTKPPSPRTCNCWP 659

Db 600 RHDRYSNRNVVFFDINMKGLDGIQGPITYGTGVFRROALYGVDAKTKPPSPRTCNCWP 659

Qy 660 KWCFCGCCFGNRKQKTKTPKTEKKLLFPKKEENQSPAYALGEIDEA--PGAENEKAG 717

Db 660 KWCFL--LCFGSRKNRKAKTVADKKK---KNREASKQIHAIENIEEGRGHKVLNVEQST 713

Qy 718 IVNQKLEKKFGQSSVFVTSITLLENGTLLKSASPASLKEAIVHISCGYEDKTDWKEIG 777

Db 714 EAMQMKLOKKYGQSPVFVASARLENGBMARNASPACLLKEAIOVISRGYEDKTEWKEIG 773

Qy 778 WIYGSVTEDILTGFMHCHGWRSIYCI PKRVAFKGSAPLNSDRLHQVLRWALGSIEIFP 837

Db 774 WIYGSVTEDILTGSKMHSHGWRHVYCTPKLAAFKGSAPINLSDRLHQVLRWALGSVEIFP 833

Qy 838 SNHCPMWYGGGLKFLERPSYINSIVPWTSLPILAYCTLPALCLTGKFTTPELNNVA 897

Db 834 SRHCPMWYGGGLKFLERLSYINSVVPWTSLPILAYCSLPAICLLTGKFTVPEISNYA 893

Qy 898 SLWMSLFCIFATSILEMNSGVCIDDMWRNEQFWIGVSSHLFAVFOGLKVIAGVD 957

Db 894 SLWFALFSSIAITGILEMGKIGIDDWMRNEQFWIGVSAHLFALFQGLKVIAGVD 953

Qy 958 TSFTVTSKGGDDEEFSELYTFKWTLLIPTTLILLNPIGVAVGVSNAINNGYESWGPFL 1017

Db 954 TNFTVTSKAADDEGFSDLYEFKWTSLIPTMTLLIINIVGIVGVSDAISNGYDSWGPFL 1013

Qy 1018 GKLFPAFWIVHLVPLFKGLVGRQNRTPPTIVIVWSILASIFSLMVRIDPFLAKDGP 1077

Db 1014 GRLEFPAFWIVHLVPLFKGLGKQDRMPTIIVWSILASITLLMVRVNPFAK-GGPI 1072

Qy 1078 LEECGGLDC 1085

Db 1073 LEICGLDC 1080

RESULT 4

T51579

cellulose synthase catalytic subunit (IRX3) - Arabidopsis thaliana

N:Alternate names: protein T10B6 80

C:Species: Arabidopsis thaliana (mouse-ear creas)

C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C:Accession: T51579

R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; M

submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51579

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1026 <SAT>

A:Cross-references: EMBL:AL391142

A:Experimental source: cultivar Columbia; BAC clone T10B6

C:Genetics:

A:Map position: 5

A:Introns: 25/3; 91/1; 184/1; 215/2; 304/2; 419/3; 507/3; 578/3; 650/1; 715/3; 833/3

A:Note: T10B6\_80

Query Match 67.0%; Score 3967.5; DB 2; Length 1026;

Best Local Similarity 66.8%; Pred. No. 6.1e-296;

Matches 736; Conservative 129; Mismatches 145; Indels 91; Gaps 16;

Qy 1 MEASAGLVAGSHNRNELVIRDDPGPKPREQNGVCOICGDDVGLAPGDPFVACNE 60

Db 1 MEASAGLVAGSHNRNELVIRHNEE--PKPLKNLDGQFCEICGDQIGLTVEGDLFVACNE 58

Qy 61 CAFPVCRDCEYERREGTQNCPOCKTRYKRLKGCQRYTGDEEEDGVDDLNEFNWDGHS 120

Db 59 CGFPACRPCEYERREGTQNCPOCKTRYKRLKGSPRVEGEDEEDIDDIYEYFNIE-HQ 117

Qy 121 --QSAVESMLYGHMSYGRG--GDPNGA-PQAFQOLNPNVPLLTNGQMVDDIPPEQHALVP 174



Db 118 DKHSHSAEAMLYGKMSYGRGPEDDENGRFP-----PVIAGHSGS----- 157  
Oy 175 SFMGGG-----KRIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAWKERMENWK 225  
Db 158 -FPVGGYGNGEHGHKRVHPY-----PSSEAGSEG---GWRERMDWK 197  
Oy 226 QROERMHOTGNDGGDDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLG 285  
Db 198 L-----QHGNLGPBPD-DDPEMGLIDEARQPLSRKVPFIASSKINPYRVIARLVILAV 250  
Oy 286 FHYRVMPVNDAFALMLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEGQ 345  
Db 251 FLRYRLNPVHDALGLMLTSVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRYERGE 310  
Oy 346 PSQAPIDFVSTVDPLKEPPLVTTNTVLSLSVDYPVDKVSCTVSDGAAMLTFEALSE 405  
Db 311 PNMALPVDVFVSTVDPLKEPPLVTSNTVLSILANDYPVEKISCTVSDGASMLTFESLSE 370  
Oy 406 TSEFAKKWVPFCCKRYNIEPRAPEMYFOQKIDYLDKVAANFVRERAMKREYEEFKVRIN 465  
Db 371 TAEFARKWVPFCCKRESIEPRAPEMYFTLKVDYLQDKVHPTEFKERRAMKREYEEFKVRIN 430  
Oy 466 ALVAKAQKVPBEGWTMODGTPWPGNNVRDHPGMIQVFLGSGGLDCEGNELPRLVYVSRE 525  
Db 431 AQVAKASKVPLEGWTMODGTPWPGNNTKDHPGMIQVFLGSHSGFDVEGHELRLVYVSRE 490  
Oy 526 KRPGYNHHKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGK 585  
Db 491 KRPGFOHHKAGAMNALVRVAGVLTNAPFMLNLDCDHYVNNSKAVREAMCFMLMDPOIGK 550  
Oy 586 VCYVQFPQRFDDGIDRHDYRANRVNVPFDINMKGLDGIQGPITYGTGCVFRRQALYGYDAP 645  
Db 551 VCYVQFPQRFDDGIDTNDYRANRNTVFFDINMKGLDGIQGPVYVGTGCVFRRQALYGYERP 610  
Oy 646 KTKKPPSRTCNCWPKWCFCCCFGNRKQKTKTEKKLLFPKKEENOSPAYALGEID 705  
Db 611 KGPKRPKMIS-----CGCCPCFGRRKKNK-----FSKNDMNGDVALG--- 649  
Oy 706 EAPGAENEKAGIVNOOKLEKKFGQSSVFVTSTLLENGTLKASAPASLKEAIVHISCG 765  
Db 650 ----GAEGDKEHLMSENNFEKTFGQSSIFVTSTLMEEGVPPSSSPAYLKEAIVHISCG 705  
Oy 766 YEDKTDMGKEIGWYGSVTEIDLTGFKMHCHGWRSTYCIPIKRVAFKGSAPNLSDRLHOV 825  
Db 706 YEDKTEWGTGLGWIYGSITEDIITGFKMCHRGWRSTYCMKRPAFKGSAPINLSDRLNQV 765  
Oy 826 LRWALGSIIEFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPTISIPLAYCTLLPAICLL 884  
Db 766 LRWALGSVEIFFSRHSPLMYGYKGGKIKWLERFAYANTTIYPTISIPLAYCILLPAICLL 825  
Oy 885 TGKFITPELNNVASLWMSLFCIPATSILEMWNGVIGIDWWRNEQFWYGGVSSHFA 944  
Db 826 TDKFIMPPISTFASLFPISLWMSIIVTIGILELWGSVISIEWWRNEQFWYGGISAHFA 885  
Oy 945 VFQGLLKVIAGVDTISFTVTSKGGDEEFSELYTFKWTLLIPTTLTLLNFIGVAVGSN 1004  
Db 886 VVQGLLKILAGIDTNTFTVTSKATDDDDFGLYAFKWTLLIPTTLTLLNFIGVAVGASD 945  
Oy 1005 AINNGYESWGPLFGKLFAPFWIVLHLYPFLKGLVGRNRTPTIVIVWSILASIFSLLMV 1064  
Db 946 AINNGYQSWGPLFGKLFPSFWIVLHLYPFLKGLMGRNRTPTIVIVWSVLLASIFSLLMV 1005  
Oy 1065 RIDPFLAKDDGPLLEECGLDC 1085  
Db 1006 RIDPFLVLTGPDTSKCGINC 1026

RESULT 5  
T05351  
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana  
N:Alternate names: protein F8B4.110  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-Jun-1999  
C/Accession: T05351  
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; D  
ewes, H.W.; Mayer, K.F.X.; Schuelker, C.  
submitted to the Protein Sequence Database, February 1999  
A/Reference number: 215409  
A/Accession: T05351  
A/Molecule type: DNA  
A/Residues: 1-1081 <BEV>  
A/Cross-references: EMBL:AL034567  
A/Experimental source: cultivar Columbia; BAC clone F8B4  
C/Genetics:  
A/Map position: 4  
A/Introns: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1;  
A/Note: F8B4.110  
C/Keywords: glycosyltransferase; hexosyltransferase; P-loop  
Query Match 66.7%; Score 3945.5; DB 2; Length 1081;  
Best Local Similarity 66.9%; Pred. No. 3.2e-294;  
Matches 727; Conservative 140; Mismatches 182; Indels 37; Gaps 11;  
Oy 1 MEASAGLVAGSHNRNELVIRDDGDPKPRPREQNGVQICGDDVGLAPGDPFVACNE 60  
Db 1 MEASAGLVAGSYRRNELVIRHSDGCTKPLKMMNGQICQICGDDVGLAETGDFVACNE 60  
Oy 61 CAFVCRDCEYERREGTQNCQCKTRYKRLKGCQRTGDEBEDGVDDLDFNFMGDHS 120  
Db 61 CAFVCRPCYEYERKDTQCCPCQCKTRFRHRGSPRVEGDEDEDVDIENEFNY----A 116  
Oy 121 QSVAESMLYGHMSYGRGDDPNGAPQAFQLPNPNVPLTNGQWVD--IPPEQHALVPSF- 176  
Db 117 QGANKA-----RHQRHGEFESSSSRHESQP-IPLLTGHTVSGEIRTPDTQSVRTTSGP 169  
Oy 177 MGGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAWKERMENWKQROER--MQGT 234  
Db 170 LGPSDRNAISSPYIDPRQVPVRIYVDPKDLNSYGLGVNDWKEVEGWMKLKQEKXNMLQMT 229  
Oy 235 GN-----DGGDDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGF 285  
Db 230 GKYNHEKGGEIEGTSGNGE--ELQMAADTRLPMRSRVPIPSRLLTPYRVVITLRLILCF 287  
Oy 286 FHYRVMPVNDAFALMLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEGQ 345  
Db 288 FLOYRTHPVKNAYPLMLTSVICEIWFAMSWILDQFPKWPPIERETYLDRLAIRYDRDGE 347  
Oy 346 PSQAPIDFVSTVDPLKEPPLVTTNTVLSLSVDYPVDKVSCTVSDGAAMLTFEALSE 405  
Db 348 PSQAPVADVFSVSTVDPLKEPPLVTANTVLSLSVDYPVDKVAQVSDGSAMLTFESLSE 407  
Oy 406 TSEFAKKWVPFCCKRYNIEPRAPEMYFOQKIDYLDKVAANFVRERAMKREYEEFKVRIN 465  
Db 408 TAEFAKKWVPFCCKFNIEPRAPEFYAQKIDYLDKQIQPSFYERERAMKREYEEFKVRIN 467  
Oy 466 ALVAKAQKVPBEGWTMODGTPWPGNNVRDHPGMIQVFLGSGGLDCEGNELPRLVYVSRE 525  
Db 468 ALVAKAQKVPBEGWTMODGTPWPGNNTKDHPGMIQVFLGSHSGLDTGNEPLRLIYVSRE 527  
Oy 526 KRPGYNHHKAGAMNALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLGK 585  
Db 528 KRPGFOHHKAGAMNALIRVSAVLTNGAYLLNVDCHYFNNSKAIKEAMCFMMDPAIGK 587  
Oy 586 VCYVQFPQRFDDGIDRHDYRANRVNVPFDINMKGLDGIQGPITYGTGCVFRRQALYGYDAP 645  
Db 588 CCYVQFPQRFDDGIDLHDYRANRNTVFFDINMKGLDGIQGPVYVGTGCCFNQALYGYDPV 647  
Oy 646 KTKKPPSRTCNCWPKWCFCCCFGNRKQKTKTEKKLLFPKKEENOSPAYALGEID 705  
Db 648 LTEE-----DLEPNIIVKSCEGSRKKGKSSKKNYERKRGV--NRSDSNAPLPMNEDID 699  
Oy 706 EAPGAENEKAGIVNOOKLEKKFGQSSVFVTSTLLENGTLKASAPASLKEAIVHISCG 765  
Db 700 EGFEGYDDERSILMSQRSVEKRFQSPVFIAATFMEQGGIPPTNPATLLKEAIVHISCG 759

[illegible]

RESULT 6  
T52054  
cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 31-Dec-2000  
C/Accession: T52054  
R/Author: T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H  
Science 279, 717-720, 1998

A:Reference number: Z13745; MUID:98111412; PMID:9445479  
A:Accession: T52054  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1065 <ARI>  
A:Cross-references: EMBL:AF027174; PIDN:AAC39336.1  
C:Genetics:  
A:Note: Ach-B  
C:Function:  
A:Description: EC 2.4.1.-; cellulose synthase [validated, MUID:98111412]; involved in as  
C:Keywords: glycosyltransferase, hexosyltransferase

Query Match	65.5%;	Score 3875.5;	DB 2;	Length 1065;
Best Local Similarity	66.7%;	Pred. No. 7.4e-289;		
Matches 728; Conservative	145;	Mismatches 161;	Indels 57;	Gaps 18;

Qy		23	DGDPGPKPPREONGOVCOICGDVDVLAPGGDPEFVACNECAFVPCRDCEYERREGTONCP	82
	:	:	:	:
Db		4	EGETAGKPMKNIVPQTCCISPDNVGKTVDGDRFVACDISFPVCRCPCYEYERKDGNQSCP	63
Qy		83	QCKTRYKRLKGCRVTGDEEEDGV-DDLDNEFNWDGHDSQSVAESMLYGHMSYGRGDPN	141
	:	:	:	:
Db		64	QCKTRYKRLKGSPAIPGDKDEDGLADEGVEFNYP--QKEKISERMIGMHLTRGK-GEEM	120
Qy		142	GAPQAFQ-LNPVNPLLTNGQMVDIIPPEQHALLVPSPMG---GGGKRTHLPY-ADPS	193
	:	:	:	:
Db		121	GEPQYDKVESHNHLPRLTSRQ--DTSGEFSSASPERLSVSSTIAGKR--LPYSSDVN	174
Qy		194	LPVQPRSMDDPSKDILAAAYGYGSVAMKERMENWKOQROER----MHOTGNDGGDDGDD---	245
	:	:	:	:
Db		175	QSPNRRIIDP-----VGLGNVAMKERVDMGMKKQEKNTGPVSTQAASERGCVDIDASTD	228
Qy		246	--ADLPLM-DEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFFHYRVMPVNDAPALM	302
	:	:	:	:
Db		229	ILADEALLNDEARQPLSRKVSIPISSRINPYRMVIMLRVILCLFLHYRTNPVENAPALM	288
Qy		303	LISVICEIWFAWSMIIDOPFKWFPIEREETYLDRSLRFDKEGOPSQLAPIDEFVSTVDPL	362
	:	:	:	:

Db	289	LVSVICEIMFALSWILDQFPKMPFVNRETYLDRLATLRDREGEPSQLAAVDIFVSTVDPL	348
Qy	363	KEPPLVTTNTVLSISVDYPVDKVSICYVSDGGAAMLTEALSETSEFAKKWVPFCRKYNI	422
Db	349	KEPPLVTANTVLSILAVDYFVDKVSICYVFDGGAAMLSFESLAETSEFARKWVPFCCKYSI	408
Qy	423	EPRAPEWYFOQKIDYLDKQVAANFVRERBRAMKREYEFEKYRINALVAKAQKVPPEGWTMQ	482
Db	409	EPRAPEWYFAKIDYLDKQVQTSFVKDRAMKREYEFEKIRINALVSKALKCPEEGWMQ	468
Qy	483	DGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNNHKKAGAMNAL	542
Db	469	DGTPWPGNNTGDHPGMIQVFLGQNGGLDABGNELPRLVYVSREKRPGFQHHKAGAMNAL	528
Qy	543	VRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDRHD	602
Db	529	VRVSAVLTNAPFILNDCDHYINNSKALREAMCFMMDPNLGQVCYVQFPQRFDGIDKND	588
Qy	603	RYANRNVVFPDINMKGLDGIQGPYVGTGCVFRQALYGYDAP--KTICKP--SRTCNC	657
Db	589	RYANRNTVFDPINLRGLDGIQGPVYVGTGCVFNRTALYGEPRPIKVHKHKKPSLSKLCG-	647
Qy	658	WPKMCFCCCCFGNRKQKKTTPKTEBKCLLFKKENOSPAYALGEIDEAAPGA--ENEX	715
Db	648	-----GSRKXNSKAKKESDKK--SGRHTDSTVVFNLDDIEEGVEGAGFDEK	694
Qy	716	AGIVNOQLEKKFGQSSVFTSTLLENGTILKSASAPSLKEAIIHVISCGYBDKTDWGE	775
Db	695	ALLMSQMSLEKRFQGSVAVFASTLMENGVPSPATPENLLKEAIIHVISCGYEDKSDWGE	754
Qy	776	IGWIYGSVTEIDLTFGMHCHGWRSIYCI PKRVAFGSAPLNLSDRLHOVLRWALGSIEI	835
Db	755	IGWIYGSVTEIDLTFGMHARGWRSIYCMPKLPFGSAPINLSDRLNQVLRWALGSVEI	814
Qy	836	FPSNHCPLMYGGGLKFLERFSYINSIYVPMWTSIPLLAYCTLPAICLLTGKFTPELAN	895
Db	815	LFSRHCPIMYGNGRLKFLERFAYVNTTIYPTISIPLMYCTLLAVCLFTNQFIIPQISN	874
Qy	896	VASLWMSLFCIPATSILEMWMSGVGDMDWRNEQWVIGVSSHLFAVFOGLKVIAG	955
Db	875	IASIMWLSLFLSIFATGILEMRWSGVGIDEMWRNEQWVIGVSAHLFAVFOGILKVIAG	934
Qy	956	VDTSFTVTSKGD--DEEPELYTEFKWTTLLIPPTLLLLNFIGVAVGSNAINNGYESWG	1014
Db	935	IDTNFTVTSKASDEDEDFAELYLFKWTTLLIPPTLLLVNLGVAVGSYAINSGYQSWG	994
Qy	1015	PLFGKLFFAFWVIVHLYPFLKGLVGRONRTPTIVIVMSILLASIFSLMVRIDPFLAKD	1074
Db	995	PLFGKLFFAFWVIVHLYPFLKGLMGRONRTPTIVIVMSVLLASISFLMVRIDPFTSRVT	1054
Qy	1075	GPLLEBGLDC 1085	
Db	1055	GPDILECGINC 1065	

RESULT 7  
 F84649  
 Probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C/Accession: F84649  
 R/Lin. X./Kaul, S./Rounsley, S.D./Shea, T.P./Benito, M.I./Town, C.D./Fujii, C.Y./  
 M./Koo, H./Moffat, K.S./Cronin, L.A./Shen, M./Vanaken, S.E./Umayam, L./Tallon, B./  
 eus, D./Nierman, W.C./White, O./Eisen, J.A./Salzberg, S.L./Fraser, C.M./Venter,  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: F84649  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-1065 <STO>  
 A/Cross-references: GB:AE002093; NID:G4432865; PIDN:AAD20713.1; GSPDB:GN00139  
 C/Genetics:







A:Molecule type: mRNA  
A/Residues: 1-685 <PEA>  
A/Cross-references: EMBL:U58284; NID:g1706957; PIDN:AA837767.1; PID:g1706958  
A/Experimental source: strain Acala S<sub>2</sub>-2; fiber  
C/Genetics:  
A/Gene: celA2  
C/Function:  
A/Description: involved in synthesis of cellulose  
C/Keywords: glycosyltransferase; hexosyltransferase

Query Match 49.1%; Score 2907; DB 2; Length 685;  
Best Local Similarity 75.0%; Pred. No. 8.9e-215;  
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

QY 411 KKWVPECKRYNIEPRAPWYFOQKIDYLKDKVAANFVRERRAMKREYEEFKVRINALVAK 470  
Db 1 RRVVPFCCKHNEPRAPPEFYFNEKIDYLDKDVHPSFVKERRAMKREYEEFKVRINALVAK 60  
QY 471 AQKVPEEGWTMODGTPWPGNNVDHPGMIQVFLGSGGLDCEGNELPRLVTVSREKRPGY 530  
Db 61 AQKPEEGWWMQDGTWPGNNTDHPGMIQVYLGAGALVDGKELPRLVTVSREKRPGY 120  
QY 531 NHHKAGAMNALVRSVAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLLCKKYCYVQ 590  
Db 121 QHHKAGAENALVRSVAVLTNAPFLNLDCDHYINNSKAMBEAMCFLMDPQFGKLCYVQ 180  
QY 591 FPRFPDGI DRHRYANRNVPFDINMKLDGIOGPIYVGTGCVFRROALYGYDAPKTKP 650  
Db 181 FPRFPDGI DRHRYANRNVPFDINMLGLDGLQGPVYVGTGCVFNROALYGYDPPVSEKR 240  
QY 651 PSRTGNCWPKMCFCCCCFGNRKOKTKTKPTEKK----KLFFKK-----EENOSP 697  
Db 241 PKMTCDWPSWC-CCCCGSGSRK--SKKKGEKGLGLGYKKKKMGKNYVKKGSAP 296  
QY 698 AYALGEIDEAAPGAEN-EKAGIYNQOKLEKKFGQSSVFVSTLLENGLTKSASPASLLK 756  
Db 297 VFPLEEIEEGLEGEYELEKSTLMSQKNFEKRFQSPVFIASTLMENGLPEGTNSTLTK 356  
QY 757 EAIHVISCGYEDKTDWGEKIGWYGSVTEDLITGFKMHCHGRSIIYCI PKRVAFKGSAPL 816  
Db 357 EAIHVISCGYEKTEWGEKIGWYGSVTEDLITGFKMHCRGWSVYCVPKRPAFKGSAPI 416  
QY 817 NLSDRHQVLRWALGSIEIFFSNHCPWYGGGLKFLERFSYINSIVYPTWSIPLLAYC 876  
Db 417 NLSDRHQVLRWALGSVEIFLSRHCPWYGGGLKWLERLAYINTIVYPTWSIPLLAYC 476  
QY 877 TLPATCLLTGKEITBELNNVASLWFMSLFICIFATSILEMWRSVGIDDMWRNEQFWYIG 936  
Db 477 TIPAVCLLTGKEITBELNNVASLWFMSLFICIFATSILEMWRSVGIDDMWRNEQFWYIG 536  
QY 937 GVSSHLPFAVFOGLLKVIAVDTSFTYTSKGGDDEEFSELTFKWTLLIPTTLLINFI 996  
Db 537 GVSAPHLFAVFOGLLKVLAGVDTNFTVITAKAADTEFGELYLFKWTLLIPTTLLINMV 596  
QY 997 GVVAGVSNAINNGYESWGPLFGKLFPAFWIVHLYPFLKGLVGRQNRPTIVIVWSILLA 1056  
Db 597 GVVAGVSDAINNGYSGWGPLFGKLFPAFWIVHLYPFLKGLMGRQNRPTIVIVWSILLA 656  
QY 1057 SIFSLLWVRIDPFLAKDGPFLBEECGLDC 1085  
Db 657 SIFSLLWVRIDPFLPKQTGPVLKQCGVEVC 685

RESULT 11  
D86157  
hypoetical protein F22D16.26 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C/Accession: D86157  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzial  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: D86157  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-1181 <STO>  
A/Cross-references: GB:AE005172; NID:g6056428; PIDN:AAF02892.1; GSPDB:GN00141  
C/Genetics:  
A/Map position: 1

Query Match 39.4%; Score 2330.5; DB 2; Length 1181;  
Best Local Similarity 43.0%; Pred. No. 3.5e-170;  
Matches 489; Conservative 166; Mismatches 274; Indels 207; Gaps 26;

QY 31 PREQGVQCICGDDVGLAPGDPFVACNECAFVPCRDCEYERREGTQNCPOCKTRYKR 90  
Db 154 PEKSGQICWLKGDCEKVVHG-----RC-ECGFRIICRDICYFDCITSGGNCPCGCKEPRYD 207  
QY 91 LKGCQRTVGDEEDGVDDLDFNEFNWMDGHSQVAESMLYGHMSYGRGDPNGAPQAFQLN 150  
Db 208 INDDPETEEDEDEDAKPL-----PQMGESKLDKRLS----- 239  
QY 151 PNVPLLTNGQVDDIPPEQHALVPSFMGGGKRIHPLPYADPSLPVQPRSMDSKDL--- 207  
Db 240 -----VKSFKAQN-----QAGDFDHTRWLFET 262  
QY 208 -AAYGYSVAMKERNENWKORQERMQTG--NDGGDDDDADLPLMDEARQOLSRKIP 263  
Db 263 KGTGYGNVWP-----KDGYGIGSGGGNGYETPPEFGERSKRLTRKVS 308  
QY 264 LPSSQINPYRMIIIRLVLGFFHYRVHRYNDAFALMLISVICEIWFAMSILDOFPK 323  
Db 309 VSAALISPYRLILALRLVALGLFTWRVRHNRREAMWLMGMSTCELMFALSWLLDQLPK 368  
QY 324 WPIERETYLDRLSLRFDK-----EGQPSQLAPIDFVSTVDPPLKEPPLVTNTVLSIL 377  
Db 369 LCPVNRITDLGLVKERFESPNLRNPKGR-SDLPGIDVFVSTADPEKEPPLVTANTILSIL 427  
QY 378 SVDYPVDKVSQVSDGAAMLTFEALSETSEFAKMWVPFCRKRYNIEPRAPWYFOQKIDY 437  
Db 428 AVDYPEKLACYLSDDGALLTFEALQTAFASTWVPFCRKHNIEPRNBEAYFGQKRN 487  
QY 438 LKDKVAANFVRERRAMKREYEEFKVRINAL-----VAKQK----- 473  
Db 488 LKNKVRIDFVRERRRVKREYDEFKVRINSLPEAIRRSDAYNVHEELRAKKQMEMMMGN 547  
QY 474 -----VPEEGWTMODGTPWPG-----NNVRDHPGMIQVFLGSGGLDCEGNE--- 515  
Db 548 NPQETVIVPKATW-MSDGSHPGTWSSGETDNSRGDHAGIIQAMLAPRNAEPVYGAADADA 606  
QY 516 -----LPRLVYSREKRPGYNNHKKAGAMNALVRSVAVLTNAPYLLNLDCHY 563  
Db 607 ENLIDTVDVDIRLPMLVYSREKRPGYNNHKKAGAMNALVRTSAIMNSGPFILNLDCHY 666  
QY 564 INNSKAIKEAMCFMMDPLLKGVYVQFPQRPDGI DRHRYANRNVPFDINMKGLDGIQ 623  
Db 667 IYNSMALREGMCFMLD-RGGDRICYVQFPQRPBGIDPNDRYANHNVTFFDYSMRALDGLQ 725  
QY 624 GPIYVGTGCVFRQALYGYDAPKTKPPSRITCNCWPKMCFCCCCFGNRKOKTKTKPTEK 683  
Db 726 GPMYVGTGCFIRITLVGFSPPRATEHHG-----W-----LGRKAYKISLR---R 767  
QY 684 KLLLFKKEENQSPAYALGEIDEAAPGAENKAGIYNQOKLEKKFGQSSVFVTS-TLLEN 742  
Db 768 PKAMMKDKDEVSLPIN--GEYNE-----EENDGDIESLLPKRFGNSNSFVASIPVAVY 820  
QY 743 GGTL-----KSASPASLL-----KEAIVISCGYEDKTDWGEKIGWYGS 782

Db 821 QGRLLIQDLQKGKNSRAGSLAVPREPLDAATVAEAIISVISC FYEDKTEWGXVGMWYGS 880  
QY 783 VTEDILTFKMHCHGMSIYICIPKRVAFKGSAPLNLSDRLHQLRWALGSIEIFPSNHCP 842  
Db 881 VTEDVVTGYRMHNRMSIYCVTKRDAFRGTAPINLTDRLHQLRWATGSEIIFSRNNA 940  
QY 843 LMYGYGGGLKFLERPSYINSIVPWTISIPLAYCTLPALCLLTGKFITPELNNVASLWFM 902  
Db 941 IF--ATRRMKFLQRAVFNVMGMPFTSLFLIVYCIIPALISLFSGQFIVQSLDITFLIYLL 998  
QY 903 SLFICIFATSIEMRWSGVIGIDMMWRNEQFWVIGVSSHLFAVFOGLKVIAGVDTSFV 962  
Db 999 SITLTLCLMLSLLEIKMSGITLHEWMRNEQFWVIGGSAHPAVALQGLKVIAGVDISFTL 1058  
QY 963 TSKG-----DDEESELTFKWTLLIPPTLLLNFIGVAVGSNAINNGYESWGPLFG 1018  
Db 1059 TSKSSAPEDGDDEFPADLYVVKMSFLMVPPLTIMVMIAIIVGLARTLYSPFPQMSKLVG 1118  
QY 1019 KLPPAFWIVVHLYPELKLGLVGRQNRPTIVIVWSILLASIFSLMWRIDPFLAKD 1074  
Db 1119 GVFFSFVWLCHLYPFAKGLMGRGRVPTIVFVWSGLSIIVSLMWYINPSPGKD 1174

RESULT 12  
T51546  
cellulose synthase catalytic subunit-like protein - Arabidopsis thaliana  
N/Alternate names: protein F2K13 60  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C/Accession: T51546  
R/Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew  
submitted to the Protein Sequence Database, August 2000  
A/Reference number: Z25394  
A/Accession: T51546  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1145 <SAT>  
A/Cross-references: EMBL:AL391141  
A/Experimental source: cultivar Columbia; BAC clone F2K13  
C/Genetics:  
A/Map position: 5  
A/Introns: 297/2; 566/3  
A/Note: F2K13\_60

Query Match 39.1%; Score 2311; DB 2; Length 1145;  
Best Local Similarity 42.6%; Pred. No. 1e-168;  
Matches 493; Conservative 170; Mismatches 271; Indels 222; Gaps 33;  
QY 11 SHNRNELVIRRDGPBPKEPREONGQVCQICGDDVGLAPG--GDPFVACNECAFVPCRD 68  
Db 100 SNTRAHLMHKVIETEPNHPQWAGSKSSCAIPGCDAKWMSDERGQDLPC-ECDFKICRD 158  
QY 69 CYEYEREGTQNCPOCKTRYKRLKGCQRVGTDEEEDGVDDLDFNEFNWDGHSQVAESML 128  
Db 159 CFIDAVKTGGGICPGCKEPYKN---THLTDQVDENG----- 191  
QY 129 YGHMSYGRGDPNGAPQAFQLPNVPLLTNGQWVDIPEQHALVPSFMGGG---KRI 184  
Db 192 -----QORPMLP---GGGGSKMERRL 209  
QY 185 HPLPYADPSLPVQPRSMDS-----KDLAAYGYSVAWKERMENWKQDERMHQTGNDG 238  
Db 210 SMVKSTNKSALMRSGTGFHDNRWLFETGTGYGNAFWT-----KDGD FG 255  
QY 239 GGDDGD-DAD-----LPLMEARQOLSRKIPLSSQINPYRMIIIRLVVLGFFHYRV 291  
Db 256 SGKDDGDGDGMGEAQDLMSRPWRPLTRKLIIPAGVISPYRLIFIRIVLALFLTWRV 315  
QY 292 MHPVNDAPALWLISVICEIWFAMSWILDQPKWFPPIERETYLRLSLRFD--KEGQ--- 346  
Db 316 KHQNPDAVWLWMSVYVCELWALSWLLDQLPKLCPINRATDLQVLKEKFEPTASNP TGK 375  
QY 347 SQLAPIDFVSVTDPLKEPPLVTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSET 406

Db 376 SDLPGFDFVSTADPEKEPPLVTANTLISILAAYPVEKLSQVSDDGALLTFEAMAEA 435  
QY 407 SEFAKKWVPCCKRYNIEPRAPENYFOQKIDYLDKDVAAANFVRERARMKREYEBFKVRINA 466  
Db 436 ASFANINWFCRKAIEPRNPDSYPSLKRDPYKNKVKSDFYDORRRVYKREFDEPKVRVNS 495  
QY 467 L-----VAKAQ-----KVPEEGWTQDGTMPGNNVR--- 493  
Db 496 LPDSIRRSDAYHAREIKAMKQQRNDEDEPMEPVKIPKATW-MADGTHMPGTWLT SAS 554  
QY 494 -----DHPGMIQVFL-----GQSGG-LDCEGNE--LPLRVYSREKRPGYNHKKKA 536  
Db 555 DHAKGDHAGIIQVMLKPPSDEPLHGVSEGFLDITDVIDRLPLRVYSREKRPGYDHKKKA 614  
QY 537 GAMNALVRSYAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLLKKVCYVQFPQRFD 596  
Db 615 GAMNALVRSALMSNGPFI LNLDCDHYIYNSEALREGMCFMMD-RGGDRLCYVQFPQRF 673  
QY 597 GIDRHRVARNVVFEDINMKGLDGIQGPITYGTGCVFRQALYGVDA PKTK-PPSRTC 655  
Db 674 GIDPSDRYANHTVFPDVNMRALDGLMGPVYVGTGCLFRRIALYGFNPFRSKDFSPS--- 730  
QY 656 NCWPKWCFCCCFGNRKQKTKPTKEKKLLFPKKEENQSPAYALGEIDEAPGAENK 715  
Db 731 -CW-----SCCF-----PSKKKNI---PEENR--ALRMSDYDDBE----- 760  
QY 716 AGIVNQKLEKKFGQSSVFVTS-----TLLENG---GTL---KSASPASL 755  
Db 761 ---MNLIVPKKFGNSTFLIDSIYVAEFGQRLADHPAVKNGRPPGALTI PRELDASTV 817  
QY 756 KEAIVHISGVEDKTDWGEIGWIGSVTEDILTFKMHCHGMSIYICIPKRVAFKGSAP 815  
Db 818 AEAIAVISGWYEDKTEWGSRIGWIGSVTEDVVTGYRMHNRGKSVYCVTKRDAFRGTAP 877  
QY 816 LNLSDRLHOVLRWALGSIEIFPSNHCPLMYGGGLKFLERPSYINSIVPWTISIPLAY 875  
Db 878 INLTDRLHOVLRWATGSEIIFPSRNALL--ASSKKILQRIAYLVNGIYPTISFLIVY 935  
QY 876 CTLPAICLLTGKFITPELNNVASLWMSLFCIFATSIEMRWSGVIGIDMMWRNEQFWVI 935  
Db 936 CFLPALSLFSGQFIVQTLNVTFLVYLLIISITLCLALLEIKMSGISLEWMRNEQFWLI 995  
QY 936 GGVSSHLEAVFOGLKVIAGVDTSFVTSK-GGD--DEESELTFKWTLLIPPTLLL 992  
Db 996 GGTSAHLAVALQGLLKVAVGEISFTLTSKSGDDIDDEFPADLYMWKTSIMIPITIM 1055  
QY 993 LNFIGNAVGSNAINNGYESWGPLFGKLFPAFWIVVHLYPELKLGLVGRQNRPTIVYWS 1052  
Db 1056 VNLIAIIVGFSRTIYSVVPQWSKLIGVFFSFVLAHLVPAKGLMGRGRRTPTIVYWS 1115  
QY 1053 ILLASIFSLMWRIDP 1068  
Db 1116 GLVAITISLLWVAINP 1131

RESULT 13  
T05646  
hypothetical protein F20D10.310 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999  
C/Accession: T05646  
R/Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K  
submitted to the Protein Sequence Database, February 1999  
A/Reference number: Z15420  
A/Accession: T05646  
A/Molecule type: DNA  
A/Residues: 1-1111 <BEV>  
A/Cross-references: EMBL:AL035538  
A/Experimental source: cultivar Columbia; BAC clone F20D10  
C/Genetics:  
A/Map position: 4  
A/Introns: 139/2; 675/3





Db 605 YGFNPFRANEYSG-----VFGQ-----EKAPAMHVRTQSQAQTS 639  
Qy 700 ALGEID-EAPGAENEKAGIVNOQKLEKFGQSSVFVTSTL-----LENG 743  
Db 640 QASDLESDTQPLNDPDLG-----LPKKEGNSMTF-TDTIPVAEYQGRPLADHMSYKNG 692  
Qy 744 ---GTLKSASP---ASLLKEATHYISCGYEDKTDWKEIGWYGSVTEIDILTGFKMHG 797  
Db 693 RPPGALLLRPPLDAPTAEAIAVISCWYEDNTWGDRIGWYGSVTEDEVVTGYRMHNRG 752  
Qy 798 WRSIYCI PKRVA PKGSAPLNSDRLHQVLRWALGSIETPSNHCPLWYGGGGLKFLERF 857  
Db 753 WRSVYCITKRDAFRGTAPINLTDRLHQVLRWATGSVEIFPSKNAMF--ATRLKFLQRY 810  
Qy 858 SYINSIVPWT SIPPLAYCTLPALCLTGKFTPELNNVASLWFMSLFICIFAT---SI 913  
Db 811 AYLNVGITPFTSIFLVVYCFPLALCLFSGKFIVQSLD---IHFLSYLLCITVTLTLISL 866  
Qy 914 LEMRSGVGIDDMWRNEQFWVIGVSSH.LFAVFQGLKVIAGVDTSTFTVTSK-GGDDEE- 971  
Db 867 LEVWMSGIGLEEMWRNEQFWLIGTSAHLAAVQGLKVIAGIEISFTLTSKASGEDED 926  
Qy 972 -FSELTEFKWTTLLIPPTLLLNFIQVAVGSNAINNGYESWGPLFGKLFPAFWIVHL 1030  
Db 927 IPADLYIVKWTGLFIMPLTIILVLAIVIGASRTIYSVTPQWGLMGGIFPSLWVLTLM 986  
Qy 1031 YPFLKGLVGRQNRPTIVIVMSILASIFSLMVRIDP 1068  
Db 987 YPFAKGLMGRGKVPITIVYVWSGLVSIWVSLWITISP 1024

RESULT 15  
T02209  
cellulose synthase (EC 2.4.1.-) catalytic chain - rice (fragment)  
C/Species: Oryza sativa (rice)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 11-Jun-1999  
C/Accession: T02209  
R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H  
Science 279, 717-720, 1998  
A/Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.  
A/Reference number: Z13745; MUID:9811412; PMID:9445479  
A/Accession: T02209  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-583 <ARI>  
A/Cross-references: EMBL:AF030052; NID:g2781432; PID:g2781433  
A/Experimental source: subsp. Japonica, strain Nipponbare  
C/Keywords: glycosyltransferase; hexosyltransferase

Query Match 35.8%; Score 2118.5; DB 2; Length 583;  
Best Local Similarity 66.2%; Pred. No. 2.4e-154;  
Matches 398; Conservative 75; Mismatches 93; Indels 35; Gaps 8;

Qy 1 MEASAGLVAGSHNRNELVIRRDGD--PGPKPREQNGOVQICGDDVGLAPGDPPVAC 58  
Db 1 MAANAGMVAAGSRNRNEFMIRPDGAPPPAKPGKSVNGQVCQICGDTVGVSATGDVFVAC 60  
Qy 59 NECAFVGRDCEYERREGTONCPQCKTRYKRLKGCQRVTGDEEEDVDLDNEFNWDGH 118  
Db 61 NECAFVCRPCYEYERKEGNQCCPQCKTRYKHKGSPRVQDEEEDVDLDNEFNK-H 119  
Qy 119 DSQSVAESMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVND---DIPPEQHALVP 174  
Db 120 GNGKGPEWQI-----QROGEDVDLSSSSSRHEQHRIPLRTSGQOISGEIIPDASPD RHSI-- 172  
Qy 175 SFMGGGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAKERMENWKQROER-MHQ 233  
Db 173 -----RSGTSSYVDPSVPVPRIVDPSKDLNSYGINSVDQERVASWRNKQDKNMQ 224  
Qy 234 TGN-----DGGDDGDADALPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVL 283  
Db 225 VANKYPARGGDMEGTGSNGE--DIQWVDARLPLSRIVPIPSNQNLNLYRIVIIIRLITL 282

Qy 284 GFFPHYRMHPVNDALALWLISVICEIWFAWSWILDQFPKWPPIERETYLDRLSLPDKE 343  
Db 283 MFFQYRVTHPVRDAGLMLVSVICEIWFAWSWILDQFPKWPPIERETYLDRALRYDRE 342  
Qy 344 GQPSQAPIDFVSTVDPLKEPPLVTTNTVLSISVDYPVDKVSQVSDGAMLTFEAL 403  
Db 343 GEPQSLAPIDVFSTVDPLKEPPLITANTVLSILAVDYVVDKVSQVSDGAMLTFEAL 402  
Qy 404 SETSEFAKQWVPFCRKNIEBPAPWYFOOKIDYLDKQVAANFVRERRAMKREYEEFKVR 463  
Db 403 SETAEFARKWVPFCCKGNIEBPAPWYFAQKIDYLDKQIPSFVKERRAMKREYEEFKVR 462  
Qy 464 INALVAKAQVPEEGWTMODGTPWPGNNVRDHPGMIQVPLGQSGGLDCBGNELPRLVYS 523  
Db 463 INALVAKAQVPEEGWTMADGTAMPGNPNRDPGMIQVFLGHSGLDTCNEIPRLVYS 522  
Qy 524 REKRPGYNHKKAGANALVRSVAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLL 583  
Db 523 REKRPGFQHHKKAGANMALIRVSAVLTTNGAYLLNVDCDHYFNNSKALREAMCFMMDPALG 582  
Qy 584 K 584  
Db 583 R 583

Search completed: December 15, 2003, 13:04:59  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:57:19 ; Search time 18 Seconds  
(without alignments)  
2837.278 Million cell updates/sec

Title: US-09-720-383C-10  
Perfect score: 5918  
Sequence: 1 MEASAGLVAGSHNRNLVLI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252	4.3	754	1	BCA1_ACEXY P19449 acetobacter
2	248	4.2	874	1	BCSA_SALTI Q82291 salmonella
3	246	4.2	874	1	BCSA_SALTY Q931n2 salmonella
4	242	4.1	739	1	BCSA_PSEFL P58931 pseudomonas
5	239.5	4.0	1596	1	ACS2_ACEXY Q59167 acetobacter
6	239	4.0	872	1	BCSA_ECO57 Q8x517 escherichia
7	238	4.0	872	1	BCSA_ECOLI P37653 escherichia
8	237	4.0	756	1	BCA2_ACEXY P82859 acetobacter
9	230.5	3.9	729	1	BCSA_XANAC P58932 xanthomonas
10	218	3.7	1518	1	BCA4_ACEXY Q9rbj2 acetobacter
11	218	3.7	1518	1	BCAS_ACEXY Q9wx75 acetobacter
12	201	3.4	1550	1	ACSI_ACEXY P21877 acetobacter
13	194	3.3	745	1	BCA3_ACEXY Q9wx61 acetobacter
14	139.5	2.4	906	1	CBL_HUMAN P22681 homo sapien
15	116	2.0	473	1	GLNA_SYNP2 P28605 synechococc
16	113.5	1.9	1254	1	POLS_EEVEE P36330 venezuelan
17	113	1.9	567	1	COX1_EMENTI P05674 venezuelan
18	113	1.9	1254	1	POLS_EEVV8 P05674 venezuelan
19	110	1.9	655	1	YS44_CAEEL Q09372 caenorhabdi
20	110	1.9	1912	1	CHD4_HUMAN Q14839 homo sapien
21	109	1.8	535	1	COX1_HANWT P48868 hansenuela w
22	109	1.8	896	1	CBL_MOUSE P22682 mus musculu
23	108.5	1.8	357	1	G6PT_HUMAN P35575 homo sapien
24	108.5	1.8	1612	1	DNM1_PARLI Q27746 paracentrot
25	108	1.8	441	1	YCDQ_ECOLI P75905 escherichia
26	108	1.8	775	1	ICP0_HSV11 P08393 herpes simp
27	107	1.8	1019	1	TOP1_USTMA P41511 ustilago ma
28	106.5	1.8	992	1	UVRA_MICLU P13567 micrococcus
29	106.5	1.8	1184	1	FBP2_HUMAN P98095 homo sapien
30	106	1.8	556	1	SRM1_SCHPO O13286 schizosach
31	106	1.8	1254	1	POLS_EEVT P09592 venezuelan
32	105.5	1.8	470	1	ROCC_BACSU P39636 bacillus su
33	105.5	1.8	565	1	YD08_MYCPN P75472 mycoplasma

34	105	1.8	873	1	COX1_ACACA Q37370 acanthamoeb
35	105	1.8	1255	1	POLS_EEYV3 P36329 venezuelan
36	104.5	1.8	524	1	COX1_WHEAT P08741 triticum ae
37	104	1.8	537	1	COX1_SCHPO P07657 schizosach
38	103.5	1.7	322	1	NUIM_STRPV P15548 strongyloce
39	103.5	1.7	524	1	COX1_ORYSA P14578 oryza sativ
40	103.5	1.7	528	1	COX1_MAIZE P08742 zea mays (m
41	103.5	1.7	530	1	COX1_SORBI P05502 sorghum bic
42	103.5	1.7	1120	1	KEFA_ECOLI P77338 escherichia
43	103.5	1.7	1332	1	XKDO_BACSU P54334 bacillus su
44	103.5	1.7	4548	1	APOA_HUMAN P08519 homo sapien
45	103	1.7	677	1	UBF1_XENLA P25979 xenopus lae

ALIGNMENTS

RESULT 1  
BCA1\_ACEXY STANDARD; PRT; 754 AA.  
ID BCA1\_ACEXY  
AC P19449;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSA.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Acetobacteraceae; Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-42 AND 196-206.  
RC STRAIN=1306-3;  
RX MEDLINE=91045951; PubMed=2146681;  
RA Wong H.C., Fear A.L., Calhoun R.D., Eichinger G.H., Mayer R.,  
RA Amikam D., Benzman M., Gelfand D.H., Meade J.H., Emerick A.W.,  
RA Bruner R., Ben-Bassat A., Tal R.;  
RT "Genetic organization of the cellulose synthase operon in Acetobacter  
RT xylinum.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8130-8134(1990).  
RN [2]  
RP ENZYME REGULATION.  
RX STRAIN=1306-3;  
RX MEDLINE=21194671; PubMed=11297407;  
RA Chang A.L., Tuckerman J.R., Gonzalez G., Mayer R., Weinhouse H.,  
RA Volman G., Amikam D., Benzman M., Gilles-Gonzalez M.-A.;  
RT "Phosphodiesterase A1, a regulator of cellulose synthesis in  
RT Acetobacter xylinum, is a heme-based sensor.";  
RL Biochemistry 40:3420-3426(2001).  
CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose in a processive way.  
CC The thick cellulosic mats generated by this enzyme probably  
CC provide a specialized protective environment to the bacterium.  
CC -1- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP  
CC + {(1,4)-beta-D-glucosyl}(N+1).  
CC -1- COFACTOR: Magnesium.  
CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
CC (c-di-GMP).  
CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -1- INDUCTION: Cellulose is produced at a linear rate with respect to  
CC cell growth when O(2) is present.  
CC -1- DOMAIN: There are two conserved domains in the globular part of  
CC the catalytic subunit: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glucosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
-----



```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37202; AAA21884.1; -.
DR PIR; A43735; A43735.
DR InterPro; IPR003919; Cellsynth_A.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF00535; Glycos_transf_2; 1.
DR PRINTS; PR01439; CELLSNTHASEA.
KW Cellulose biosynthesis; Transferrase; Glycosyltransferase;
KW Transmembrane; Inner membrane.
FT DOMAIN 147 240 CATALYTIC SUBDOMAIN A.
FT DOMAIN 317 377 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 167 187 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 430 450 POTENTIAL.
FT TRANSMEM 468 488 POTENTIAL.
FT TRANSMEM 516 536 POTENTIAL.
FT TRANSMEM 549 569 POTENTIAL.
FT ACT_SITE 189 189 POTENTIAL.
FT ACT_SITE 333 333 POTENTIAL.
FT BINDING 236 236 SUBSTRATE (POTENTIAL).
FT BINDING 238 238 SUBSTRATE (POTENTIAL).
SQ SEQUENCE 754 AA; 84442 MW; 8D5FC1FE62E2C068 CRC64;

Query Match 4.3%; Score 252; DB 1; Length 754;
Best Local Similarity 20.4%; Pred. No. 6.3e-11;
Matches 168; Conservative 89; Mismatches 217; Indels 350; Gaps 36;

QY 275 IITRLVVGFFPHYRVMPVNDALWLISVCEIWFMSWILDQ---FPKWP----- 326
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 53 LIVALVCTLVFFIVGRGKSRRTQIF-LEVLASLVSLRY-LTWRLTETLDFDTWIGGIGV 110
QY 327 --IEREY-LDRLSLRFDKEGQPSQIAP-----IDFFVSTVDPLKEPPLVTN 371
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 111 TLLMAELIYALYMLFLSYFQTIQPLHRAPLPLPDNDWDPTVDIFIPYD--EQLSIYRL 167
QY 372 TVLSILSVDPVVDKVSQVSDGAMLTFEALSETSEFAKKVVPFCRXYNIERAPENYF 431
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 168 TVLGALGIDWPPDKVNVYILDDG----- 190
QY 432 QOKIDYLDKVAANFVRERAMKEEYEFKVRINALVAKAKVPEEGWTMDGTPWPGNN 491
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 191 -----VRPEFEQF-----AK----- 200
QY 492 VRDHPGMIOVFLGSGGLDCEGNEELRLVYVSREKRPQYVNHKKAGAMALVRVSAVLTN 551
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 201 -----DCGA-----LYIGRVD---SSHAKAGNLNHAIK---RTS 228
QY 552 APYLLNLDCHYINNSKAIKEAMCFMMDPLLGKKVCYVQPFQDFDIDRHRVYANRVYF 611
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 229 GDYILLDGDH-IPTRAFQIAMGWM---VADRKIALMQTPHHFVSPD----- 272
QY 612 FDINMGKLDGIQGPITYGTGCVFRQALYGYDAPKTKKPSRTCNCWPKWCFCCCFGNR 671
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 273 -----PQGNLAVGY-----RTPPE-----GN- 289
QY 672 KOKKTKPKTEKGLLFKKKEENGSPAYALGEIDEAPGAEENEKAGIVNOQKLEKKFGQS 731
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 290 -----LFY-----GVIOD---GNDFDA 304
QY 732 SVFVTSLLLENGTLKSASPASLKEAIVHISCGYEDKTDWGKEIGMITYGVTEIDILTF 791
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 305 TFF-----CGSCALIRREAIESIG-----GFAVETVTEDAHTAL 338
```

```
QY 792 KMCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIIEFSPNHCPLWYGGG 850
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 339 RMQRRGWSTAYLRIP---VASGLATERLTHIGQRMWRARGMIQIFRYDNPML-----GGG 391
QY 851 LKFLERFSYINSIYVWTSIPLAYCTLPACILLTGKFI--TPELNVVASLWFMSLFIC 907
   ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 392 LKLGRLCYLSAMTSFFFAIPRVIFLASPLAFLFFGQNTIAASP-----LA 437
QY 908 IFATSILEMWMS---GVGIDDWWRNEQFW-VIGGVSSHLFAVFOGLKVIAGVDTSTVT 963
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 438 VLAYAIPHMFHSIATAKVNKGWR-YSEFSEVYETTMALFLVRVTITITLMPSPKGFNV 496
QY 964 SKGG--DDEEPESELYTFKWTLLIPPTILLNFIGVAGVSNAINNGYESWGPLFG--K 1019
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 497 EKGGLVEEEDF-----LGATYPNIIIFAGIMTLGLLGLFE 532
QY 1020 LFFAFWVIHLVPLPKGLVGRQNRTPPIVWMS-----ILLASI 1059
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 533 LTFHF-----NQLAGIAKRAYLNCIWMISLIILLAAI 566

RESULT 2
BCSA_SALTI STANDARD; PRT; 874 AA.
ID_BCSA_SALTI
AC Q8Z591.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR STY4181 OR T3898.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC uridine 5'-diphosphate glucose to cellulose, which is produced as
CC an extracellular component for mechanical and chemical protection
CC at the onset of the stationary phase, when the cells exhibit
CC multicellular behavior (rdar morphotype). Co-expression of
CC cellulose and thin aggregative fimbriae leads to a hydrophobic
CC network with tightly packed cells embedded in a highly inert
CC matrix (By similarity).
CC -1- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC + {(1,4)-beta-D-glucosyl}(N+1).
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC (c-di-GMP) (By similarity).
CC -1- PATHWAY: Bacterial cellulose biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
```

```
CC (potential).
CC -!- DOMAIN: There are two conserved domains in the globular part of
CC the protein: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL627281; CAD08006.1; -.
DR EMBL; AE016847; AAC071373.1; -.
DR InterPro; IPR003919; Cellynth_A.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
DR PRINTS; PR01439; CELSINTHASEA.
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane; Complete proteome.
KW DOMAIN 271 364 CATALYTIC SUBDOMAIN A.
FT DOMAIN 441 501 POTENTIAL.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 525 545 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 592 612 POTENTIAL.
FT TRANSMEM 634 654 POTENTIAL.
FT TRANSMEM 668 688 POTENTIAL.
FT TRANSMEM 833 853 POTENTIAL.
FT ACT_SITE 313 313 POTENTIAL.
FT ACT_SITE 457 457 POTENTIAL.
FT BINDING 360 360 SUBSTRATE (POTENTIAL).
FT BINDING 362 362 SUBSTRATE (POTENTIAL).
SQ SEQUENCE 874 AA; 100020 MM; 3395D09CD0F51990 CRC64;

Query Match 4.2%; Score 248; DB 1; Length 874;
Best Local Similarity 18.0%; Pred. No. 1.6e-10;
Matches 152; Conservative 100; Mismatches 198; Indels 394; Gaps 32;

QY 258 LSRKIPLPSSQINPYRMIIIRLVVGFHFHYVMHPVN--DAFAL-----W 302
DB 189 LVRRMPGRFSAL----MLIVLSLTVSCRYIMWRYTSTLNWDDPVSLVCGILLLFAETYAW 244
QY 303 LISVCEIWFAMSWILDQFPKWFPIERTYLDRLSLRFDKEGQPSQAPIDFFVSTVDPL 362
DB 245 IVLVLG--YFQVWPLNRQP--VPLPKEM-----SQWPTVDIFVPTYN-- 283
QY 363 KEPPLVTTNTVLSILSDYVPVDKVCVSDGAAMLTFEALSETSEFAKKWVPCKRYNI 422
DB 284 -EDLNWVKNTIYASLIGIDWPKDKLTINILDDG----- 315
QY 423 EPRAPWYFQOKIDYLDKVAANFVRERRAMKREYEEKVRINALVAKAQKVPBEGWTMQ 482
DB 316 -----RESFRQF----- 322
QY 483 DGTWPNGNNVRDHPGMIQVFLGSGGLDCGNELRLVYVSREKRPQVNHKKAGAMNAL 542
DB 323 -----ARHVGVA-----YIAR-----ATHEHAKAGNINNA 347
QY 543 VRVSAVLTNAPYLNLDCDHYINNSKAIKEAMCMMDPLLGKVCVYQFPQRPDGDID--- 599
DB 348 LKHA-----KGEFVAIFDCDHVPTRSFLQMTGWFLKE-----KQLAMQTFHHFFSPDPFE 399
```

```
QY 600 ----RHDRYANRNVVFPDINMKGLDGIQGPITYVTGCVFRQALGYDAPKTKPPSRTC 655
DB 400 RNLGRFRKTPNEGTLFGLVQDND----- 424
QY 656 NCWPKWCFCCCFGNRKQKTKTPTEKKLLFFKKEENQBPAYALGEIDEAPGAENEK 715
DB 425 -MWDATPFCGSC----- 435
QY 716 AGIVNQKLEKKEFGQSSVFVTSTLLENGTLKSASPASLKEAIVHISCGYBDKTDWKE 775
DB 436 -AVIRKPLD-----E 445
QY 776 IGWI-YGSVTEDILTFKMHCHGWSIYC-IPKRVAFKGSAPLNLSDRLHQLRWALGSI 833
DB 446 IGGIAVETVTEDAHTSLRLHRRGYTSAYMRIPOSA--GLATESLSAHIGQIRWARGMV 502
QY 834 EIPFSNHPDLWYGGGLKFLERFSYINSIVYPTWSIPLAYCTLPALCLTGKFTPEL 893
DB 503 QIFRLDN-PL--FGKGLKLAQRLCYLNAMFHLGSLPRLIFLTAFLALHAYI----- 554
QY 894 NNVASLWMSLFICIFA-----TSILEMRWSGVGIDWNRNEQF-----WVIGVSSHL 942
DB 555 -----IYAPALMIALFVIPHWVHASLNSKIQGKYRHSFW-SEIYETVLAWYIAPT--- 605
QY 943 FAVFQGLIKYIAGVDTSFVTWSKGDDEEFSELYTFKWTLLI--PPTLLLLNFIGVVA 1000
DB 606 -----LVALINPHKGNKFNVTAKGLVEE-----KYVDWISRPYIFVLNLNLGVAA 652
QY 1001 GVSNAINGYESWGPLFGKLFPAFWIVHLYPFLKGLVGRONRPTIIVWSILASIFS 1060
DB 653 GV-----WRYTYGP-----ENETLTVIV-----S 671
QY 1061 LLWV 1064
DB 672 LVWV 675

RESULT 3
BCSA_SALTY
ID_BCSA_SALTY STANDARD; PRT; 874 AA.
AC 093IN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR STM3619.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14028;
RX MEDLINE=21160181; PubMed=11260463;
RA Zogaj X., Nimtz M., Rohde M., Bokranz W., Roemling U.;
RT "The multicellular morphotypes of Salmonella typhimurium and
RT Escherichia coli produce cellulose as the second component of the
RT extracellular matrix.";
RL Mol. Microbiol. 39:1452-1463 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21927388; PubMed=11929533;
RA Solano C., Garcia B., Valle J., Betasain C., Ghigo J.-M., Gamazo C.,
RA Iasa I.;
RT "Genetic analysis of Salmonella enteritidis biofilm formation:
RT critical role for cellulose.";
RL Mol. Microbiol. 43:793-808 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spleth J., Clifton S.W., Latreille P.,
```

RA Courtney L., Porwollk S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
LT2.";  
RL Nature 413:852-856(2001).  
CC -I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose, which is produced as  
CC an extracellular component for mechanical and chemical protection  
CC at the onset of the stationary phase, when the cells exhibit  
CC multicellular behavior (rdar morphotype). Co-expression of  
CC cellulose and thin aggregative fimbriae leads to a hydrophobic  
CC network with tightly packed cells embedded in a highly inert  
CC matrix.  
CC -I- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP  
CC + {(1,4)-beta-D-glucosyl}(N+1).  
CC -I- COFACTOR: Magnesium (By similarity).  
CC -I- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
CC (c-di-GMP) (By similarity).  
CC -I- PATHWAY: Bacterial cellulose biosynthesis.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -I- DOMAIN: There are two conserved domains in the globular part of  
CC the protein: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ315770; CAC44015.1; -.  
DR EMBL; AE008867; AAL22479.1; -.  
DR EMBL; AJ315148; AAC86199.1; -.  
DR StyGene; SG27272; bcsA.  
DR InterPro; IPR003919; CellSynth\_A.  
DR InterPro; IPR005150; Cellulose\_synt.  
DR InterPro; IPR001173; Glyco\_trans\_2.  
DR Pfam; PF03552; Cellulose\_synt; 1.  
DR Pfam; PF00535; Glycos\_transf\_2; 1.  
DR PRINTS; PR01439; CELLSYNTHASEA.  
DR Cellulose biosynthesis; Transferase; Glycosyltransferase;  
KW Transmembrane; Inner membrane; Complete proteome.  
FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.  
FT DOMAIN 441 501 CATALYTIC SUBDOMAIN B.  
FT TRANSMEM 30 50 POTENTIAL.  
FT TRANSMEM 151 171 POTENTIAL.  
FT TRANSMEM 173 193 POTENTIAL.  
FT TRANSMEM 230 250 POTENTIAL.  
FT TRANSMEM 525 545 POTENTIAL.  
FT TRANSMEM 547 567 POTENTIAL.  
FT TRANSMEM 592 612 POTENTIAL.  
FT TRANSMEM 634 654 POTENTIAL.  
FT TRANSMEM 668 688 POTENTIAL.  
FT TRANSMEM 833 853 POTENTIAL.  
FT ACT\_SITE 313 313 POTENTIAL.  
FT ACT\_SITE 457 457 POTENTIAL.  
FT BINDING 360 360 SUBSTRATE (POTENTIAL).  
FT BINDING 362 362 SUBSTRATE (POTENTIAL).  
SQ SEQUENCE 874 AA; 100043 MW; 4C9421B58606310A CRC64;  
Query Match 4.2%; Score 246; DB 1; Length 874;  
Best Local Similarity 18.1%; Pred. No. 2.2e-10;

Matches 153; Conservative 99; Mismatches 198; Indels 394; Gaps 32;  
QY 258 LSRKIPBSSQINPYRMIIIRLVVLGFFHRYVMHPVN--DAFAL-----W 302  
DB 189 LVRKMPGRFSAL---MLIVSLTVCSCRYIMWRYTSTLANDDPVSIVCGILLFAETVAV 244  
QY 303 LISVICBIFWFSWILDPFKPFPIRETYLDRLSIRFDKGGQPSQAPIDFFVSTVDPL 362  
DB 245 IIVLVLG--YFQVWVPLNRQP--VPLPKEM-----SQWPTVDIFVPTYN-- 283  
QY 363 KEPPLVTNTVLSILSVDPYVDKVCYVSDGGAAMLTFEALSETSEFAKKWVPCKRYNI 422  
DB 284 -EDLVNVTNTIYASIGIDWPCKXINWILDGGR----- 316  
QY 423 EPRAPEMYFOQKIDYLDKQVAANFVERRAMKREYEEFKVRINALVAKQKVPEEGWTMQ 482  
DB 317 -----ESFR----- 320  
QY 483 DGTWPNGNVRDHPGMIOVFLGGSGLDCEGNELPRLVYVSREKRPYNNHKKAGAMNAL 542  
DB 321 -----HFARHVG-----VHYIARTT---HEHAKAGNINNA 347  
QY 543 VRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKVCYVQPFQPDGID-- 599  
DB 348 LKHA---KGEFVAIFDCDHVPTRSFQMTGWFLKE---KQAMQTPHHFFSPDPFE 399  
QY 600 ---RHDRYANRNVFFEDINMKLDGIQGPITYGTGCVFRQALYGYDAPKTKPPSRIC 655  
DB 400 RNLGRFRKPTNEGTLFYGLVQDND----- 424  
QY 656 NCWPKWCFCCCFGNRKQKTKTKTEKKLLFFKKEENQSPAYALGEIDEAPGAENEX 715  
DB 425 -MMDATFFCGSC----- 435  
QY 716 AGIVNQKLEKKEFGQSSVFVTSTLLENGTLKSASPASLKEAIVHISGYEDKTDWKE 775  
DB 436 -AVIRKRPD----- 445  
QY 776 IGMV-YGVTEDILTGFKMCHGMSIYC-IPKRVAFKSGAPLNSDRLOVLRWALGSI 833  
DB 446 IGGIAVEYTEDAHTSLRLHRRGYSAYMRIPQAA--GLATESLSAHIGQIRIWRAGMV 502  
QY 834 EIFFSNHCPIMYGGGGLKFLERFSYINSIVPWTSLPLAYCTLPALCLTGKFTPEL 893  
DB 503 QIFRLDN-PL--FGKGLKLAQRLCYNAMFHLSGIPRLIFLAPLAFLLHAYI---- 554  
QY 894 NNVASLWFSWLFICIFA-----TSILEMRWSGVGIDDWWRNEQF-----WVIGVSSH 942  
DB 555 -----IYAPALMIALFVIPHVHASLNSKIQGKYRHSFW-SEIYETVLAWYIAPT--- 605  
QY 943 FAVFQGLKVIAGVDTSTVTSKGGDEEFSELYTFKWTLLI--PPTLLLLNFIGVVA 1000  
DB 606 -----LVALINPHKGFNTAKGLVEE-----KYVDVVISRPYIFLVLLNLGVAA 652  
QY 1001 GVSNAINNGYESWGPLFGKLPFAFWIVVHLPPFLKGLVGRQNRPTIIVVWSILLASIFS 1060  
DB 653 GV-----WRYYYGP-----ENETLTVIV-----S 671  
QY 1061 LLWV 1064  
DB 672 LVWV 675  
RESULT 4  
ID BCSA\_PSEFL STANDARD; PRT; 739 AA.  
AC P58931; Q8RS21;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSA OR WSSB.  
OS Pseudomonas fluorescens.



OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SBW25;  
RX MEDLINE=22013850; PubMed=12019221;  
RA Spliers A.J., Kahn S.G., Bohannon J., Travisano M., Rainey P.B.;  
RT "Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness."  
RT fitness."  
RL Genetics 161:33-46(2002).  
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes uridine 5'-diphosphate glucose to cellulose, which is produced as an extracellular component responsible for the structural integrity and rigidity of self-supporting mats characteristic of the "wrinkly spreader" phenotype.  
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl} (N) = UDP + {(1,4)-beta-D-glucosyl} (N+1).  
CC -!- COFACTOR: Magnesium (By similarity).  
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid (c-di-GMP) (By similarity).  
CC -!- PATHWAY: Bacterial cellulose biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).  
CC -!- DOMAIN: There are two conserved domains in the globular part of the protein: the N-terminal domain (domain A) contains the conserved DXD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the QXXRW motif and is present only in processive glycosyl transferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AY074776; AAL71842.1; -.  
DR InterPro: IPR001173; Glyco\_trans\_2.  
DR Pfam: PF00535; Glycos\_transf\_2; 1.  
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
KW Transmembrane; Inner membrane.  
FT DOMAIN 157 250 CATALYTIC SUBDOMAIN A.  
FT DOMAIN 327 387 CATALYTIC SUBDOMAIN B.  
FT TRANSMEM 36 55 POTENTIAL.  
FT TRANSMEM 59 76 POTENTIAL.  
FT TRANSMEM 83 101 POTENTIAL.  
FT TRANSMEM 116 138 POTENTIAL.  
FT TRANSMEM 417 436 POTENTIAL.  
FT TRANSMEM 440 462 POTENTIAL.  
FT TRANSMEM 524 546 POTENTIAL.  
FT TRANSMEM 551 573 POTENTIAL.  
FT ACT\_SITE 199 199 POTENTIAL.  
FT ACT\_SITE 343 343 POTENTIAL.  
FT BINDING 246 246 SUBSTRATE (POTENTIAL).  
FT BINDING 248 248 SUBSTRATE (POTENTIAL).  
SQ SEQUENCE 739 AA; 82165 MW; 2B962EA3854B23BB CRC64;  
Query Match 4.1%; Score 242; DB 1; Length 739;  
Best Local Similarity 19.1%; Pred. No. 3.5e-10;  
Matches 155; Conservative 86; Mismatches 207; Indels 364; Gaps 33;  
OY 260 RKIPLPSSQINPYRMIIIRLVV-----IGF-----FFHYRVMHPVN 296  
DB 77 RKIP-----GRLAILALVVLGLVASLRVMFWRLTSTLGFTWVDMFFGYGLV---- 123  
OY 297 DAFALWLISVIGEIMFAMSWIILDDQFKMFPFIERETYLDRLSLRFDEKQPSQLAPIDFFV 356

DB 124 -AAEFYALIVLIFGYVQTAMPRLRTPVWLKTE-----PEEWPTVDVFI 165  
OY 357 STVDPLEKPPLVTTNTVLSISVDYPVDKVCVSDGAAMLTFEALSETSEPAKKVVP 416  
DB 166 PTYN---EALSIKLTTPAAQAMDWPKDLRVHVLDDG-----RRDDFRE----F 208  
OY 417 CKRYNIEPRAPPEWYFOQKIDYLDKKVAANFVRERRAMKREYEEFKRINALVAKQVPE 476  
DB 209 CR-----KGVNYIR----- 218  
OY 477 EGVWTQDGTMPGNNVRDPGMIVQLGSGGLDCEGNELPRLVYVSREKRPQVNHKKA 536  
DB 219 -----RDN-----NFAKA 227  
OY 537 GAMNALVRVSAVLTPNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCYVQFQRF- 595  
DB 228 GNINEALKV---TDGEYIALFDADHVPTRSFLQVSLGWFLKDP---KLAMLOTTHFFF 279  
OY 596 -----DGIDRHRYANRVVFPDINMKGLDGIGPIYVGTGCVFRQALYGYDAPTKK 649  
DB 280 SPDPPEKNLDTFRAVPNEGELFYGLVODGND----- 310  
OY 650 PPSRTNCWPKWCFCCCEGFRNRKQKTTKPKTEKKLLFPKKEENQSPAYALGEIDEAAP 709  
DB 311 -----LWNATFFC----- 318  
OY 710 GAENEKAGIVNOQLEKKEGQSSVFTSTLLENGGTLKSASPASLLKEAIVHISGEYEDK 769  
DB 319 -----GSCAVIRREPLLEIGV-----AVE----- 338  
OY 770 TDMGKEIGWYGSVTEDILTGFKMHCHGWRSTY-CIPKRVAFKGSAPLNSDRLHOYLRW 828  
DB 339 -----TVTEDAHTALKLNRLGYNTAYLAIPQAA---GLATESLSRHINDRIRW 383  
OY 829 ALGSIIEIFPSMCPPLWYGGGLKPLERPSYINSIVYFWTSIPLAYCTLPAICLTGKF 888  
DB 384 ARGMAQIFRTDN-PL--LGKGLKWGORICYANAMQHFFYGLPRLVFLTAPLAYLIFG-- 437  
OY 889 ITPELNVAASLWFMSLFI-CIFATSILEMWMSGVGIDWRNEQF-----WVIGVSSH 942  
DB 438 --AEIFHASALMIVAVVLPPLVHSSLTNSRIQGRFRHSFW-NEVETVLAWYI-----L 488  
OY 943 FAVFOGLLKVIAGVDTSPFVTSKGC-DDEFSSELYTFKWTLLIPTLLLNFIGVAVG 1001  
DB 489 PPVLVALVNPKAG--GFNVTDKGIIIDQF-----FDW-KLARPYLVLLAVNLIGLGRG 539  
OY 1002 VSNAINNGYESWGPLFGKLFPAF---WVIYHL 1030  
DB 540 IHQLI-----WGDASTAVTVAINLTWTLTYNL 565  
RESULT 5  
AC2\_ACEXY  
ID AC2\_ACEXY STANDARD; PRT; 1596 AA.  
AC Q59167;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase 2 [Includes: Cellulose synthase catalytic subunit (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose synthase 2 regulatory domain)].  
DE AC2A11.  
GN AC2A11.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Acetobacteraceae; Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 23769;  
RX MEDLINE=95394846; PubMed=7665515;  
RA Saxena I.M., Brown R.M. Jr.;  
RT "Identification of a second cellulose synthase gene (acsA11) in

RT Acetobacter xylinum.";  
RL J. Bacteriol. 177:5276-5283(1995).  
CC -I- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP  
CC + {(1,4)-beta-D-glucosyl}(N+1).  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -I- DOMAIN: There are two conserved domains in the globular part of  
CC the catalytic subunit: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -I- MISCELLANEOUS: It is not essential for cellulose production in  
CC this strain.  
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
CC GLYCOSYLTRANSFERASE FAMILY 2.  
CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U15957; AAA85264.1; -  
DR PIR: T31338; T31338.  
DR InterPro: IPR003919; CellSynth\_A.  
DR InterPro: IPR003920; CellSynth\_B.  
DR InterPro: IPR001173; Glyco\_trans\_2.  
DR Pfam: PF03170; BCSB; 1.  
DR Pfam: PF00535; Glycos\_transf\_2; 1.  
DR PRINTS; PR01439; CELLSINTHASEA.  
DR PRINTS; PR01440; CELLSINTHASEB.  
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
KW Transmembrane; Inner membrane.  
FT DOMAIN 1 749 CATALYTIC.  
FT DOMAIN 750 1596 CYCLIC DI-GMP BINDING (POTENTIAL).  
FT DOMAIN 145 238 CATALYTIC SUBDOMAIN A.  
FT DOMAIN 315 375 CATALYTIC SUBDOMAIN B.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 106 126 POTENTIAL.  
FT TRANSMEM 396 416 POTENTIAL.  
FT TRANSMEM 421 441 POTENTIAL.  
FT TRANSMEM 505 525 POTENTIAL.  
FT TRANSMEM 544 564 POTENTIAL.  
FT TRANSMEM 1553 1573 POTENTIAL.  
FT ACT\_SITE 187 187 POTENTIAL.  
FT ACT\_SITE 331 331 POTENTIAL.  
FT BINDING 234 234 SUBSTRATE (POTENTIAL).  
FT BINDING 236 236 SUBSTRATE (POTENTIAL).  
SQ SEQUENCE 1596 AA; 175799 MW; D98A6F6259E1F3CE CRC64;  
  
Query Match 4.0%; Score 239.5; DB 1; Length 1596;  
Best Local Similarity 18.2%; Pred. NO. 1.6e-09;  
Matches 145; Conservative 91; Mismatches 228; Indels 331; Gaps 26;  
  
QY 274 MIIIRLVLGFFHYRVMPVNDALFALWLIS-----VICEIWFAMSWILDQFQWFP 326  
Db 81 LLVSLRYMWMRLTTTLEHSPLOALSLLLVAELVALTLCLSYFQMSWPLDRKP--LP 138  
QY 327 IEREYTLRLSLRFDKEGQSLAPIDFVSTVDPLKEPPLVTNTVLSILSVDPYDKV 386  
Db 139 LPADT-----TDWEVDVYVPSYN---EELSLVRSTVLGALAIDWPKXL 180  
QY 387 SCYVSDGAMLTFEALSETSEFAKKWVPFCKRYNIEPRAPDEWYFQOKIDYLKDKVAANF 446  
Db 181 NVYILDDG----- 188

QY 447 VRERRAMKEEYEFKVRINALVAKQKVEEGWTMODGTPWPGNNVRDHPGMIQVFLGQS 506  
Db 189 -----RRSFHAFAMEAGA-----GYIIRDQ----- 209  
QY 507 GGLDCEGNELRLVYVSREKRPGVNHHKAGAMNALVRVSAYLTNAPYLLNLDCHYINN 566  
Db 210 -----NNHAKAGNLNHALRV-----TEGEYVIFDCDH-IPT 240  
QY 567 SKAIKAMCFMM-DPLGKKVCYVQFQRFDGIDRHR--YANRV-----VFEDINMKG 618  
Db 241 RGFLLKKTIGMMADP---KLALQTPHHFYSPPFQRLNLTGQNVPEGNMFYGLVODG 296  
QY 619 LDGIQPIYVGTGCVFRQALYGYADPKTKPPSRTCNCWPKWCFCGCCFGNRKOKTKTK 678  
Db 297 NDFWDATPFCGSCAIRSAVLGIG----- 322  
QY 679 PKTEKKLLFPKKEENOSPAYALGEIDEAPGAENKAGIVNOQKLEKFGQSSVFVTST 738  
Db 323 -----FATET 327  
QY 739 LLENGTLKSASPASLLKEAIVHISGVEDKTDWGKEIGWYGVTEIDILTGFKMHCHGW 798  
Db 328 -----VTEDAH TALKMQRG 343  
QY 799 RSIYCIKRAVAFKGSAPLNLSDRLHQLVRLWALGSIEIFFSNHCPLVYGGGLKFLERFS 858  
Db 344 HTAYL--RQPLAAGLSTERLMHIGQVRWARGMQLQIMRLDN-PL--LGSGLRWQRLC 397  
QY 859 YINSIVYPTWISPLAYCTLPALICLTGKFTPELNNVASLWFMSL-----FICIFAT 911  
Db 398 YLSAMSHFLPAIPRLVFLASPLAFLFCQ-----NIIASPPAILVYAFPHVFHSIGTL 451  
QY 912 SILEMRWGVGIDDWNRNQFW-VIGGVSSHFAVFOGLLKVIAGVDTSTFTVTSKG- 969  
Db 452 SRVEGRW-----RYSFWSSEIYETTLALFLVRVTIMTLNPKGEFNVTDKGLLQ 501  
QY 970 EEFSELYTFKWTLLIPTTLTLNFIQVAVGSNAINNNGYESWGPLFGKLFFAFWVIYH 1029  
Db 502 SEYFDL-----NAVYPNVILAVILALVLRIGIGMWE-YHDLRLAQSFALNTLWAVS 554  
QY 1030 LYPFLKGL-VGRQNR 1043  
Db 555 LIIVLASIAVGRETR 569  
  
RESULT 6  
BCSA\_ECO57 STANDARD; PRT; 872 AA.  
ID BCSA\_ECO57  
AC Q8X5L7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSA OR Z4948 OR ECS4413.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; Pubmed=11206551;  
RA Berna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck B.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;

```

RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC uridine 5'-diphosphate glucose to cellulose, which is produced as
CC an extracellular component for mechanical and chemical protection
CC at the onset of the stationary phase, when the cells exhibit
CC multicellular behavior (rod morphology). Co-expression of
CC cellulose and thin aggregative fimbriae leads to a hydrophobic
CC network with tightly packed cells embedded in a highly inert
CC matrix (By similarity).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC + {(1,4)-beta-D-glucosyl}(N+1).
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC (c-di-GMP) (By similarity).
CC -!- PATHWAY: Bacterial cellulose biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- DOMAIN: There are two conserved domains in the globular part of
CC the protein: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005579; AAG58675.1; ALT_INIT.
DR EMBL; AP002565; BAB37836.1; ALT_INIT.
DR InterPro; IPR003919; Celsynth_A.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
DR PRINTS; PR01439; CELLSNTHASEA.
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane; Complete proteome.
FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.
FT DOMAIN 441 501 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 525 545 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 592 612 POTENTIAL.
FT TRANSMEM 640 660 POTENTIAL.
FT TRANSMEM 668 688 POTENTIAL.
FT TRANSMEM 833 853 POTENTIAL.
FT ACT_SITE 313 313 POTENTIAL.
FT ACT_SITE 457 457 POTENTIAL.
FT BINDING 360 360 SUBSTRATE (POTENTIAL).
FT BINDING 362 362 SUBSTRATE (POTENTIAL).
SQ SEQUENCE 872 AA; 99710 MW; F3F1A24A2B713FBA CRC64;
Query Match 4.0%; Score 239; DB 1; Length 872;
Best Local Similarity 17.9%; Pred. No. 7.5e-10;
Matches 147; Conservative 98; Mismatches 198; Indels 380; Gaps 33;
QY 274 MILLRLVVLGFFHYRVMHPVDAFALW--LISVICEL-----WFAMSWILDQFP 322

```

```

Db 201 MLIVLSLTVSCRITWRYTSTLN-----WDDPSLVCGLLILFAETVAMIVL--VLGYFQ 253
QY 323 KMFPIERETYLDRLSLRFDEKQSPQLAPIDPFVSTVDPLKEPPLVTTNTVLSLSVDYP 382
Db 254 VVWPLNRQP---VPLPKDMSLWPS---VDIFVPTYN--EDLNVTAKNTIYASLSGIDWP 302
QY 383 VDKVSCYVSDGAAMLTFEALSETSEFAKKVVPFCRRYNIERAPAEWYFOOKIDYLDKV 442
Db 303 KDKLINIWLDGQ----- 315
QY 443 AANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWTMODGTPWPGNNVRDHPGMIQVF 502
Db 316 -----REEFRQF-----AQNV----- 326
QY 503 LGQSGGLDCEGNELPRLVYVSREKRGYNHKKAGAMNALVRVSAVLTNAPYLLNLDCH 562
Db 327 -----GVKTIARTT---HEHAKAGINNALKYA---KGEFVSIFDCDH 363
QY 563 YINNSKAIKEAMCPMDPDLGKKVCYVQFPQRFQID-----RHDRYANRNVVFPDIN 615
Db 364 VPTRSFLQMTVGWFLKE---KQLAMQTPHHFPSPDPFERNLGRFRKTPNEGTLFYGLV 419
QY 616 MKGLDGIQGPITYGTCVFRQALYGYDAPKTKPPSRTCNCWPKWCFCCECGNRKQKK 675
Db 420 QDQND----- 435
QY 676 TTKPKTEKKLLFPKKEENQSPAYALGEIDEAPGAENEKAGIVNQKLEKKFGQSSVFV 735
Db 436 ----- 444
QY 736 TSTLLENGTLKSAPASLLKEAIVHISCGVEDKTDWGEIEMI-YGSVTEIDILTFKMH 794
Db 445 -----EIGGIAVETVTEDAHTSLRLH 465
QY 795 CHGWRSIYC-IPKRVAFKGSAPLNSDRLHQVLRWALGSIEIFSNHCPLMYGGGLKP 853
Db 466 RRGYTSAYMRIPQAA---GLATESLSAHIGQIRIRWARGMVQIFRLDN-PL---TGKGLKF 518
QY 854 LERFSYINSIVYPWTSIPLLAYCTLPALCLTGKFTPELNNVASLWMSLFCIFPA--- 910
Db 519 AQLCYVNAHFHFLSGIPRLIFLTAPLAFLLHAYI-----IYAPALMALFVLPH 569
QY 911 ---TSILEMRWSGVGIDDWWRNEQF---WVIGCVSSHFAVFGGLKLVAGVDTSTFTV 962
Db 570 MIHASLTNSKIQKVRHSFW-SEIYETVLAWYIAPPT-----LVALLNPHKGFENV 619
QY 963 TSKGG-DDEEFSELYTFKWTLLIPTTLLLNFGVAVGVSAINNGYESWGLFGKLF 1021
Db 620 TAKGGLVEEYVD---W-VISRPIFLVLNLVGAVGI-----WRYFYG--- 660
QY 1022 FAFWVIVHLYPFLKGLVGRNRTPTIVIVMSILASIFSLMV 1064
Db 661 -----PPT-----EMLTVVVSMVWV 675
RESULT 7
BCSA_ECOLI STANDARD; PRT; 872 AA.
ID BCSA_ECOLI
AC P37653; P37654; P76712; P76713; Q8RSS7;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR B3533.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
MEDLINE=94316500; PubMed=8041620;

```



RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
RT "Analysis of the *Escherichia coli* genome. V. DNA sequence of the  
RT region from 76.0 to 81.5 minutes.";  
RL Nucleic Acids Res. 22:2576-2586(1994).  
RN [2]  
RP REVISIONS.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of *Escherichia coli* K-12.";  
RL Science 277:1453-1474(1997).  
RN [3]  
RP CHARACTERIZATION.  
RC STRAIN=ECOR10, ECOR12, and TOB1;  
RX MEDLINE=21160181; PubMed=11260463;  
RA Zogaj X., Nimtz M., Rohde M., Bokranz W., Roemling U.;  
RT "The multicellular morphotypes of *Salmonella typhimurium* and  
RT *Escherichia coli* produce cellulose as the second component of the  
RT extracellular matrix.";  
RL Mol. Microbiol. 39:1452-1463(2001).  
CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose, which is produced as  
CC an extracellular component for mechanical and chemical protection  
CC at the onset of the stationary phase, when the cells exhibit  
CC multicellular behavior (rdar morphotype). Co-expression of  
CC cellulose and thin aggregative fimbriae leads to a hydrophobic  
CC network with tightly packed cells embedded in a highly inert  
CC matrix.  
CC -1- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP  
CC + {(1,4)-beta-D-glucosyl}(N+1).  
CC -1- COFACTOR: Magnesium (By similarity).  
CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
CC (c-di-GMP).  
CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -1- DOMAIN: There are two conserved domains in the globular part of  
CC the protein: the N-terminal domain (domain A) contains the  
CC conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC OXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -1- MISCELLANEOUS: The genes bcsA, bcsB, bcsZ and bcsC are  
CC constitutively transcribed but cellulose synthesis occurs only  
CC when adra, a putative transmembrane protein regulated by agfD, is  
CC expressed. Cellulose production is abolished in E.coli K12.  
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
CC frameshift in position 128.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U00039; AAB18510.1; ALT\_FRAME.  
DR EMBL; U00039; AAB18511.1; ALT\_FRAME.  
DR EMBL; AE000430; AAC76558.1; ALT\_INIT.  
DR EcoGene; EG12260; bcsA.  
DR InterPro; IPR003919; Cellsynth\_A.  
DR InterPro; IPR001173; Glyco\_trans\_2.  
DR Pfam; PF00535; Glycos\_transf\_2; 1.  
DR PRINTS; PR01439; CELLSYNTHASEA.  
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
KW Transmembrane; Inner membrane; Complete proteome.

FT	DOMAIN	271	364	CATALYTIC SUBDOMAIN A.
FT	DOMAIN	441	501	CATALYTIC SUBDOMAIN B.
FT	TRANSMEM	30	50	POTENTIAL.
FT	TRANSMEM	151	171	POTENTIAL.
FT	TRANSMEM	173	193	POTENTIAL.
FT	TRANSMEM	230	250	POTENTIAL.
FT	TRANSMEM	525	545	POTENTIAL.
FT	TRANSMEM	547	567	POTENTIAL.
FT	TRANSMEM	592	612	POTENTIAL.
FT	TRANSMEM	640	660	POTENTIAL.
FT	TRANSMEM	668	688	POTENTIAL.
FT	TRANSMEM	833	853	POTENTIAL.
FT	ACT_SITE	313	313	POTENTIAL.
FT	ACT_SITE	457	457	POTENTIAL.
FT	BINDING	360	360	SUBSTRATE (POTENTIAL).
FT	BINDING	362	362	SUBSTRATE (POTENTIAL).
SO	SEQUENCE	872 AA;	99784 MW;	143268A2EB228F7 CRC64;

Query Match 4.0%; Score 238; DB 1; Length 872;  
Best Local Similarity 17.9%; Pred. No. 8.9e-10;  
Matches 147; Conservative 98; Mismatches 198; Indels 380; Gaps 33;

QY	274	MIILRLVVLGFFHYRVHMPVNDALW--LISVICEI-----WFAMSWILDQFP	322
Db	201	MLIVLSLTVSCRYIMWRYTSTLN-----WDDPSLVCGLLFLFAETVAVIVL--VLGYFQ	253
QY	323	KMFPIERETVLDRLSLRFDKQGQPSQAPLIDFVSTVDPLKEPPLVTNTVLSLSVDYP	382
Db	254	VVWPLNRQP---VPLPKDMSLWPS---VDIFVPTYN--EDLNVVKNTIYASLGIQWP	302
QY	383	VDKVSQVSDGAAMLTFEALSETSEPAKRVVPFCKRYNIEPAPERYFOQKIDYKQKV	442
Db	303	KDKLNIWILDGG-----	315
QY	443	AANFRRERAMKREYEEFKVRINALVAKQVPEEGWTMDGTPWPGNNVRDHPGMIQVF	502
Db	316	-----REFRQF-----AQNV-----	326
QY	503	LGQSGGLDCEGNELPRLVYVSREKRPYVNHKKAGAMNALVRVSAVLTNAPYLLNDGDH	562
Db	327	-----GVKIYART--HEHAKAGINNALKYA---KGEFVSIFDCDH	363
QY	563	YINNSKAIKEAMCFMMDPLLKKVCYVQFQRFIDID-----RHDRYANRVVFFDIN	615
Db	364	VPTRSFLQMTMGWFLKE---KQLAMQTPHHFFSPDPFERNLGRFRKTPNEGTLFYGLV	419
QY	616	MKGLDGIQPIYVGTGCVFRQALGYDAPRTKKPPSRTCNCWPKWCFCGCCFGNRKQKK	675
Db	420	QDQND-----	435
QY	676	TTKPKTEKKLLPFKKEENQSPAYALGEIDEAAPGAENEKAGIVNQKLEKKFGQSSVFV	735
Db	436	-----AVIRRKPLD-----	444
QY	736	TSTLLENGTLKSPASLLEKAIHVISGVEDKTDWKEIGWI-YGSVTEDILLTGFKMH	794
Db	445	-----EIGGIAVETVEDAHTSLRLH	465
QY	795	CHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQLRWALGSIEIFFSNHCPLRYGGGLKF	853
Db	466	RRGYTSAYMRIPQAA---GLATESLSAHIGQRIRWARGMVQIFRLDN-PL---TGKGLKF	518
QY	854	LERFSYINSIVPWTISIPLLAYCTLPALCLTGKFTPELNNVASLWPMSLFICIFA---	910
Db	519	AQRLCYVNAMFHLFSGIPRLIFLTAFLALLHAYI-----IYAPALMIALFVLPH	569
QY	911	---TSILEMRWSGVGIDDWNRNQF---WIGGVSSHLPVFOGLKLYIAGVDTSTFTV	962
Db	570	MIHASLTSNKIQGKYRHSFW-SEIYETVLAWYIAPT-----LVALINPHKGFNV	619
QY	963	TSKGG-DDEEFSELVTFKWTLLIPPTLLILNFIGVAVGSNAINNGVSGPLFGKLF	1021
Db	620	TAKGGLVEEYVD---W-VISRPYIFLVNLVGVAVGI-----WRYFYG---	660

Qy 1022 FAFWVIVHLYPFLKGLVGRNRPPTIVIVWSILLASIFSLMW 1064  
Db 661 -----pPT-----EMLTVVSMWV 675

RESULT 8  
BCA2\_ACEXY STANDARD; PRT; 756 AA.

AC 082859;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSA.  
OC Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OX Acetobacteraceae; Gluconacetobacter.  
RN NCBI\_TaxID=28448;  
RP [1]  
RC SEQUENCE FROM N.A.  
RC STRAIN=BPR 2001;  
RX MEDLINE=98296257; PubMed=9630539;  
RA Nakai T., Moriya A., Tonouchi N., Tsuchida T., Yoshinaga F.,  
RA Horinouchi S., Sone Y., Mori H., Sakai F., Hayashi T.;  
RT "Control of expression by the cellulose synthase (bcsA) promoter  
RT region from Acetobacter xylinum BPR 2001.";  
RL Gene 213:93-100(1998).  
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose. The thick cellulosic  
CC mats generated by this enzyme probably provide a specialized  
CC protective environment to the bacterium (By similarity).  
CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-beta-D-glucosyl] (N) = UDP  
CC + [(1,4)-beta-D-glucosyl] (N+1).  
CC -!- COFACTOR: Magnesium (By similarity).  
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
CC (c-di-GMP).  
CC -!- PATHWAY: Bacterial cellulose biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -!- DOMAIN: There are two conserved domains in the globular part of  
CC the catalytic subunit: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB010645; BAA31463.1; -.  
DR InterPro; IPR003919; CellSynth\_A.  
DR InterPro; IPR001173; GlycoTrans\_2.  
DR Pfam; PF00535; GlycosTransf\_2; 1.  
DR PRINTS; PR01439; CELLSTHASEA.  
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
KW Transmembrane; Inner membrane.  
FT DOMAIN 147 242 CATALYTIC SUBDOMAIN A.  
FT DOMAIN 319 379 CATALYTIC SUBDOMAIN B.  
FT TRANSMEM 27 47 POTENTIAL.  
FT TRANSMEM 49 69 POTENTIAL.  
FT TRANSMEM 106 126 POTENTIAL.  
FT TRANSMEM 167 187 POTENTIAL.  
FT TRANSMEM 409 429 POTENTIAL.  
FT TRANSMEM 432 452 POTENTIAL.

FT TRANSMEM 470 490 POTENTIAL.  
FT TRANSMEM 517 537 POTENTIAL.  
FT TRANSMEM 551 571 POTENTIAL.  
FT ACT\_SITE 189 189 POTENTIAL.  
FT ACT\_SITE 335 335 POTENTIAL.  
FT BINDING 238 238 SUBSTRATE (POTENTIAL).  
FT BINDING 240 240 SUBSTRATE (POTENTIAL).  
SQ SEQUENCE 756 AA; 84561 MW; 6954F39A25E73B0A CRC64;

Query Match 4.0%; Score 237; DB 1; Length 756;  
Best Local Similarity 19.2%; Pred. No. 8.6e-10;  
Matches 155; Conservative 96; Mismatches 247; Indels 308; Gaps 31;

Qy 269 INPYRMIIIRLVVLGFPPHYRVMHPVNDAFALWLSVCEIWFAMSWILDQ---FPKWF 325  
Db 47 LNNNEQLIVAAVCVIFVVGGRKSRRTQIF-LEVLASLVSLRY-LTWRLTETLDFTW 104  
Qy 326 P-----IERETV-LDRLSLRFDEKQPSQLAP-----IDFVSTVDPLKEP 365  
Db 105 QGILGVILLMAELVLMFLSYFQTIQPLHRAPLPLPDNDWDPTVDIFIPTD---EQ 161  
Qy 366 PLVTTNTVLSVSDYFVDKVCYVSDDGAAMLTFEALSETSEFAKWPVFCRNVIEPR 425  
Db 162 LSVRLTVLGAIGIDWPPDKVNYIILDDG----- 190  
Qy 426 APEWTFQOKIDYLDKVAANFVRRRAKREYEEFKVINALVAKAQKVPBEGWTMDGT 485  
Db 191 -----VRPEEQFAKDCGALYI----- 207  
Qy 486 PWPENNVRDHPGMIVQLGSGGLDCEGNELPLVYVSREKRPGYNHKKAGAMNALVRV 545  
Db 208 -----GRVDVD-----SAHAKAGNLHAIK- 227  
Qy 546 SAVLTNAPYLLNDCHYINNSKAIKEAMCFMMDPLGKVCYVQFPQRFQDIDRHRYA 605  
Db 228 --RTSGDYILILDCH-IPTRAFLQIAMGM--VADRKIALMQTPHFFYSPD----- 274  
Qy 606 NRVNVFPDINMKGLDIGPIYVGTGCVFRQALYGDAPKTKKPSRTCNCWPKWCFC 665  
Db 275 -----PQRNLAVGY-----RTPPE----- 289  
Qy 666 CCFGRKOKTKTPKTEKKLLFPKKEENQSPAYALGEIDEAAPGAENKAGIVNQKLE 725  
Db 290 --GN-----LFY-----GVIOD--G 300  
Qy 726 KKEGSSVFVTSTLLENGTLKSASPASLKEAIVHISGEYEDKTDWKEIGVYSYTE 785  
Db 301 NDFWDATFF-----CGSCALIRREAIESIG-----GFAVEYTE 334  
Qy 786 DILTFKWHCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHOVLWALGSIIEFFSNHCP 844  
Db 335 DAHTALRMQRGWSSTAYLRIP---VASGLATERLTTHIGQRMWRARGMQIFRVDPNML- 390  
Qy 845 YGVGGGLKFLERFSYINSIVPWTSLPILAYCTLPALCTLTGKFITPBLNVASLWPM 904  
Db 391 ---GRGLKGLQRLCYLSAMTSFFFAIPRVIFLASPLAFLFAGQNIITAAPLAVAAALPH 447  
Qy 905 FICIPATSIEMRWSGVGIDWWRNEQFW-VIGGVSSHLPFAVQGLKVIAGVDTSEFTT 963  
Db 448 MFHSIATA-----AKVNGWR-YSFVSEVYETTMALFLVRVTIVTLLFPSKGFENV 498  
Qy 964 SKGG--DDEEFSELYTFK--WTLTLLPPTLLLNLFIGVAVGSNAIINGESWGPLFG 1018  
Db 499 EKGGVLEEEEFDLGATYPIIIFATIMGGLLIGLFELLYRFNQLDVIARNAY----- 550  
Qy 1019 KLFPAFWIVHLYPFLKGL-VGRQNR 1043  
Db 551 -LINCAMALLISLILFAAIVGRETK 575  
RESULT 9  
BCSA\_XANAC STANDARD; PRT; 729 AA.  
ID\_BCSA\_XANAC

```

AC P58932;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR XAC3518.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_taxid=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Farnsworth E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
CC uridine 5',-diphosphate glucose to cellulose, which is produced as
CC an extracellular component for mechanical and chemical protection
CC (By similarity).
CC -I- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC + {(1,4)-beta-D-glucosyl}(N+1).
CC -I- COFACTOR: Magnesium (By similarity).
CC -I- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
CC (c-di-GMP) (By similarity).
CC -I- PATHWAY: Bacterial cellulose biosynthesis.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -I- DOMAIN: There are two conserved domains in the globular part of
CC the protein: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE012000; AAM38361.1; -
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane; Complete proteome.
FT DOMAIN 151 244 CATALYTIC SUBDOMAIN A.
FT TRANSMEM 321 381 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 405 425 POTENTIAL.

```

FT	TRANSMEM	427	447	POTENTIAL.
FT	TRANSMEM	520	540	POTENTIAL.
FT	TRANSMEM	549	569	POTENTIAL.
FT	TRANSMEM	610	630	POTENTIAL.
FT	ACT_SITE	193	193	POTENTIAL.
FT	ACT_SITE	337	337	POTENTIAL.
FT	BINDING	240	240	SUBSTRATE (POTENTIAL).
FT	BINDING	242	242	SUBSTRATE (POTENTIAL).
SQ	SEQUENCE	729 AA;	80915 MW;	B9C08BB995E795B1 CRC64;
Query Match				
Best Local Similarity		3.9%;	Score 230.5;	DB 1; Length 729;
Matches 139;		Conservative 82;	Mismatches 185;	Indels 331; Gaps 26
QY	299	FALWLISVCEIMFAMSWILDQFPKPIERETYLDRLSLRFDKGQPSQAPIDFFVST	358	
DB	110	FILGLGLGALYAFILVLGYQVLPWRKP---VPLPADQLWPS---VDVFIPT	161	
QY	359	VDPLKEPPLVTTNTVLSLSVDYPVDKVSQVSDGAAMLTFEALSETSEFAKWVPFCK	418	
DB	162	YN---EPLSVRTTVLAASVIDWPAGKITIHLLDDG-----	194	
QY	419	RYNIEPRAPEWYFQOKIDYLKDKVAANFVRERBRAMKREYEEFKVRINALVAKQVBEQ	478	
DB	195	-----RR-----DEFR-----	200	
QY	479	WTMQDGTWPNGNNVRDHPGMIQVFLGSGGLDCEGNELPRLVYVSREKRPGYNHKKAGA	538	
DB	201	-----AFCAEVG-----INVTYRTN---NAHAKAGN	223	
QY	539	MNALVRVSALVTNAPYLLNLDCDHYINNASKAIKEAM-CPMMDPLGKKVCYQFPQRF--	595	
DB	224	INAALK---KCSGDYVAIFDCDH-IPTRSLQVAMGWFLHD---TKLALVQMPHYFFS	274	
QY	596	----DGDHRHRYANRNVPFEDINMKGLDGIQPIYVGTGCVPRQALYGYDAPRTKKP	650	
DB	275	PDPFERNLDTGKVPNEGELFYGLLDGNDQWNATFFCGSCAVIKRTAL-----	323	
QY	651	PSRTCNCWPKWCFCCCCFGNRKQKTTKPKTEKKLLFPKKEENQSPAYALGEIDEAPG	710	
DB	324	-----	323	
QY	711	AENEKAGIVNQKLEKKFGQSVFTSTLLENGTLKSASPASLKEAIVHISCGYEDKT	770	
DB	324	-----	323	
QY	771	DMGKEIGMI-YGSVTEDILTFGRMHCHGWRISYICIPKRVAFKGSADPLNSDRLHQVLRWA	829	
DB	324	---EEVGVAVETVEDAHTAKLQRRGYRTAYLAVPQAA--GLATESLSGHVAQRIRWA	378	
QY	830	LGSEIEIFPSNHCPLWYGGGIGKELERFSYINSIVYPWTSIPLAYCTLPAICLLTGKFI	889	
DB	379	RGMAQIARIDN-PL--LGRGKLQSRLCYLNAMLHFYGVPRITYLTAPLAYLFFGAHV	434	
QY	890	TPELNNVASLWFMSLFICIFA-TSILEMRWSGVGIDDMWR---NEQF---WVIGVS	939	
DB	435	-----IQASALMILAYALPHILLQANLTNLRVQSRFRHLLMNEVETTLAWYI----	481	
QY	940	SHLFAVQGLLKVIAGVDT-SFTVTSKGG-----DDEEFSELTFEKWTTLLIPPTLLL	992	
DB	482	-----FRPTLVALLNPKLGKENVTPKGGVARSYFDAQIAKPYLF-----LLL	524	
QY	993	LNFIGVAVGVSMAINNG 1009		
DB	525	LNVGWVAGVLRILIYVG 541		
RESULT 10				
BCA4_ACEXY				
ID	BCA4_ACEXY	STANDARD;	PRT;	1518 AA.
AC Q9RBJ2;				
28-FEB-2003 (Rel. 41, Created)				
28-FEB-2003 (Rel. 41, Last sequence update)				



DT 28-FEB-2003 (Rel. 41, last annotation update)  
DE Putative cellulose synthase 2 [includes: Cellulose synthase  
DE catalytic subunit [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding  
DE domain (Cellulose synthase 2 regulatory subunit)].  
GN BCSABI-A.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Acetobacteraceae; Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 7664 / IFO 13693;  
RX MEDLINE=99310341; PubMed=10382968;  
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,  
RA Inoue Y.;  
RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM  
RT 7664: implication of a novel set of cellulose synthase genes.";  
RL DNA Res. 6:109-115(1999).  
CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP  
CC + ((1,4)-beta-D-glucosyl)(N+1).  
CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (potential).  
CC -1- DOMAIN: There are two conserved domains in the globular part of  
CC the catalytic subunit: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
CC GLYCOSYLTRANSFERASE FAMILY 2.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB015803; BAA7593.1; -.  
DR InterPro; IPR003919; Cellsynth\_A.  
DR InterPro; IPR003920; Cellsynth\_B.  
DR InterPro; IPR001173; Glyco\_trans\_2.  
DR Pfam; PF03170; BcGB; 1.  
DR Pfam; PF00535; Glycos\_transf\_2; 1.  
DR PRINTS; PR01439; CELLSYNTHASE.  
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
KW Transmembrane; Inner membrane.  
FT DOMAIN 1 731 CATALYTIC.  
FT DOMAIN 732 1518 CYCLIC DI-GMP BINDING (BY SIMILARITY).  
FT DOMAIN 144 237 CATALYTIC SUBDOMAIN A.  
FT DOMAIN 314 374 CATALYTIC SUBDOMAIN B.  
FT TRANSMEM 24 44 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 105 125 POTENTIAL.  
FT TRANSMEM 404 424 POTENTIAL.  
FT TRANSMEM 427 447 POTENTIAL.  
FT TRANSMEM 465 485 POTENTIAL.  
FT TRANSMEM 514 534 POTENTIAL.  
FT TRANSMEM 543 563 POTENTIAL.  
FT TRANSMEM 1481 1501 POTENTIAL.  
FT ACT\_SITE 186 186 POTENTIAL.  
FT ACT\_SITE 330 330 POTENTIAL.  
FT BINDING 233 233 SUBSTRATE (POTENTIAL).  
FT BINDING 235 235 SUBSTRATE (POTENTIAL).  
SQ SEQUENCE 1518 AA; 166404 MW; 0EC99B35B6DE4543 CRC64;

Query Match 3.7%; Score 218; DB 1; Length 1518;  
Best Local Similarity 18.0%; Pred. No. 6.4e-08;  
Matches 145; Conservative 88; Mismatches 232; Indels 340; Gaps 25;  
QY 275 IIRLVLVGFFPHRYRMHPVNDALFALWL-----ISVCEIWFAMSWILDQFPKWPPIE 328  
DB 75 LMLSLVLSRYIVWRLTATVQ--FSNWLQALAVLLLAAYALMTLCLSYFQMAWPLR 132  
QY 329 RETYLDRLSLRFDEKQPSQLAPIDFVSTVDPLKEPPLVTNTVLSTLSVDYPVDKVS 388  
DB 133 RREH-----PLPEDMAQWPS-----VDVFVPSYN---BELSLVSTVLGALDLDPADRLNV 181  
QY 389 YVSDGAAMLTFEALSETSEFAKKWVPFCRKYNIERAPAEWYFQOKIDYLDKQVAANFVR 448  
DB 182 YLDDG----- 187  
QY 449 ERRAMREYESEFKVRINALVAKQKVPBEGWTMODGTPWGNVVRDHPGMIQVFLQSGG 508  
DB 188 ---RRKAFHDFAVEAGA----- 201  
QY 509 LDCEGNELPRLVYVSREKPGYNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINSG 568  
DB 202 -----GYIIRAE---NNHAKAGNLNHALAV---TDSPPAVIFDCDHPTRGF 242  
QY 569 AIKEAMCFMMDPLGKKVCYVQPFQFIDRDRYANRVYFFDINMKGLDIGPIYV 628  
DB 243 LRRTIGWMAVP---NLALLQPHHFYAPDPFQR-----NLAG--GMHVP--- 282  
QY 629 GTGCVFRQALYGYDAKTKKPSRTCNCPKWCFCGCCFGNRKQKTKTKTEKKLLF 688  
DB 283 PEGNMFYGLVODGND-----FWDATPFCGSC----- 308  
QY 689 FKEENSPAYALGEIDEAPGAENEKAGIVNOQLEKKFGQSSVFTSTLLENGTLKS 748  
DB 309 ----- 308  
QY 749 ASPASLKEAIIHIVISGVEDTKWKGIGWYGSVTEDILTFKMHCHGRSICYIPKRV 808  
DB 309 ---AIRREAVMGIG-----GFATETVTEDAHALKMQRGWTAYLREPLA 352  
QY 809 AFKGSAPLNLSDRLHQLRWALGSIIEFFSNHCPLMWYGGLKFLERFSYINSIVYPT 868  
DB 353 A--GLATERLLIHIGQVRWARGMQIMRLDNPM-----GAGLRWEQRLCYLSAMSHFLF 406  
QY 869 SIPLAAYCTLPALICLTGKFTPELNNVASLWMF-SLFCIFATSIEMRWSGVGIDDMW 927  
DB 407 AIPRLTFLVSPFLAFLGQNTIASPLAISVYALPHIFHSVTLSRIEGRW----- 457  
QY 928 RNEQFM-VIGGVSSHLFAVFOGLKVIAGVDTSPFTVTSKGDDEEFSSELYTFKWTLLIP 986  
DB 458 -RYSFWESEIYETSLALFLVRITVTLQPHKGFNTDKG----- 497  
QY 987 PTTLLNFIGVAVGSNAINNGYESWGPLFGKLFFAFWIVHLVYPLKGLVGRQNRTP 1046  
DB 498 -----LLARGYFDWDVAVPNVILAGVLCAL---LRGVFG----- 529  
QY 1047 IVIWMGI-----LLASIFSLWVRI 1066  
DB 530 --IWMQFHDRLALQSFILNTLWVI 552  
RESULT 11  
BCAS\_ACEXY STANDARD; PRT; 1518 AA.  
ID BCAS\_ACEXY  
AC Q9WX75;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, last sequence update)  
DT 28-FEB-2003 (Rel. 41, last annotation update)  
DE Putative cellulose synthase 3 [includes: Cellulose synthase  
DE catalytic subunit [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding  
DE domain (Cellulose synthase 3 regulatory subunit)].  
GN BCSABI-B.  
OS Acetobacter xylinus.

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 7664 / IFO 13693;
RX MEDLINE=99310341; PubMed=10382968;
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA Inoue Y.;
RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT 7664: implication of a novel set of cellulose synthase genes.";
RL DNA Res. 6:109-115(1999).
CC -1- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
CC + {(1,4)-beta-D-glucosyl}(N+1).
CC -1- PATHWAY: Bacterial cellulose biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- DOMAIN: There are two conserved domains in the globular part of
CC the catalytic subunit: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC GLYCOSYLTRANSFERASE FAMILY 2.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AB015804; BAA77600.1; -.
DR InterPro; IPR003919; Celleynth_A.
DR InterPro; IPR003920; Celleynth_B.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF031170; BCSB; 1.
DR Pfam; PF00535; Glycos_transf_2; 1.
DR PRINTS; PR01439; CELLSNTHASEA.
DR PRINTS; PR01440; CELLSNTHASEB.
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane.
FT DOMAIN 1 731 CATALYTIC.
FT DOMAIN 732 1518 CYCLIC DI-GMP BINDING (BY SIMILARITY).
FT DOMAIN 144 237 CATALYTIC SUBDOMAIN A.
FT DOMAIN 314 374 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 105 125 POTENTIAL.
FT TRANSMEM 404 424 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 465 485 POTENTIAL.
FT TRANSMEM 514 534 POTENTIAL.
FT TRANSMEM 543 563 POTENTIAL.
FT TRANSMEM 1481 1501 POTENTIAL.
FT ACT_SITE 186 186 POTENTIAL.
FT ACT_SITE 330 330 POTENTIAL.
FT SITE 233 233 IMPORTANT FOR SUBSTRATE BINDING
FT SITE 235 235 (POTENTIAL).
FT SITE 235 235 IMPORTANT FOR SUBSTRATE BINDING
FT SITE 235 235 (POTENTIAL).
SQ SEQUENCE 1518 AA; 166464 MW; 7D7634503183DAB6 CRC64;
Query Match 3.7%; Score 218; DB 1; Length 1518;
Best Local Similarity 18.0%; Pred. No. 6,4e-08;
Matches 145; Conservative 88; Mismatches 232; Indels 340; Gaps 25;

```

QY		275	IIIRLVLTGFFPHRVMPVNDAFALWL-----ISVICEIWFAMSWILDOFPKWPPIE	328
Db		75	LMLSLVLSTLRYIVWRLTATVQ--FSNWLQTALAVALLLAEAYALMTLCISYFQMAMPLR	132
QY		329	RETYLDRLSLRFDEKGQPSQLAPIDFVSVTDPLKEPPLVTNTVLISVDYPVKVSC	388
Db		133	RREH----PLPEDMAQWPS---VDVFVPVSYN---ELSLVRSTVLGALDLMDPADRLNV	181
QY		389	YVSDDGAAMLTFEALSETSEFAKKWVPFCRKRYNIERAPAWFYQQIKIDYLKDVAANFVR	448
Db		182	YILDG-----	187
QY		449	ERRAKREYEFEFKVRINALVAKAQKVPEEGWTQDGTWPGNMNRDPGMIOVFLGSGG	508
Db		188	---RRKAHFDFAVEAGA-----	201
QY		509	LDCEGNELPRLVYVSREKRPGYNHHKAGANNALVRSAVLTNPAYLLNDCDHYNNSK	568
Db		202	-----GYIRAE---NNHAKAGNLHALAV---TDSPFVIFDCOHVPTRGF	242
QY		569	AIKEAMCFMMDDLGLKKVCYVQFPQRFDDIGDRHRDYANRNVPFDINMKGLDGIOGPITY	628
Db		243	LRTIGMMADP---NLALLQTPHHFYAPDPQR-----NLAG--GMHV---	282
QY		629	GTCGVFRROALYGVDAPKTCKPPSRTCNCWPKWCFCGCCFGNRKQKTKTKPTEKKLLF	688
Db		283	PEGNMFGYLVQDGD-----FWDATFFCGSC-----	308
QY		689	EKKEENOSPAYALGEIDEAAPGAENEKAGIVNQKLEKKFGQSSVFTSTLLENGTLKS	748
Db		309	-----	308
QY		749	ASPASLKEAIHVISCGYEDKTDWGKEIGWIYGSYTEDIITGFMCHGWRSIYCIPRV	808
Db		309	--AIRREAVMGIC-----GFATEYTEDAHALKMQRRGWGTAYLRPLA	352
QY		809	AFKGSAPLNLSRHLQVLRWALGSIEIFFSNHCPMYGGGKLFLERFSYINSIVYPWT	868
Db		353	A--GLATERLIHGQVRVARGMIQIMRLDNPM-----GAGLRWEQRLCYLSAMSHFLF	406
QY		869	SIPPLAYCTLPALCILLTGKFIITPELNNAVSLWM-SLFICIFATSILEMRWSGVGIDDWW	927
Db		407	AIPLRLFVSPFLAFLEGONITIASPLAISVYALPHIFHSVITLSRIEGRW-----	457
QY		928	RNEQFW-VIGGVSSHLEFAVEQGGLLKVIAGVDTSFVTSKGGDDEEFSELTFKMTLLIP	986
Db		458	-RYSFMSIEYETSLALEFLVRITIVTLLQPHKGFNTDKG-----	497
QY		987	PTLLLLNFIGNVAGVSNAINNGYESWGPLEGKLFPAFWIVHLYPFLKGLVGRONRTP	1046
Db		498	-----LLARGYFDWDVAVPNVILIAGVLCAL--LRGVFG-----	529
QY		1047	IVIWMSI-----LLASIFSLLMWRI	1066
Db		530	--IWQFHDLRALQSFILNTLMWVI	552

RESULT 12  
ACSI\_ACEXY STANDARD; PRT; 1550 AA.

ID ACSI\_ACEXY  
AC P21877; P37717;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Cellulose synthase 1 [includes: Cellulose synthase catalytic domain  
[UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose  
synthase 1 regulatory domain)].  
GN ACSAB OR ACSA OR ACBS.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Acetobacteraceae; Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 6-20.  
RC STRAIN=ATCC 53582;  
RX MEDLINE=91346705; PubMed=2151718;  
RA Saxena I.M., Lin F.C., Brown R.M. Jr.;  
RT "Cloning and sequencing of the cellulose synthase catalytic subunit  
RL gene of *Acetobacter xylinum*.";  
RN Plant Mol. Biol. 15:673-683(1990).  
[2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 768-781.  
RC STRAIN=ATCC 53582;  
RX MEDLINE=91322509; PubMed=1830823;  
RA Saxena I.M., Lin F.C., Brown R.M. Jr.;  
RT "Identification of a new gene in an operon for cellulose biosynthesis  
RL in *Acetobacter xylinum*.";  
RN Plant Mol. Biol. 16:947-954(1991).  
[3]  
RP REVISIONS.  
RC STRAIN=ATCC 53582;  
RX MEDLINE=94364954; PubMed=8083166;  
RA Saxena I.M., Kudlicka K., Okuda K., Brown R.M. Jr.;  
RT "Characterization of genes in the cellulose-synthesizing operon (acs  
RL operon) of *Acetobacter xylinum*: implications for cellulose  
RN crystallization.";  
RP J. Bacteriol. 176:5735-5752(1994).  
[4]  
RP SEQUENCE OF 1-8 FROM N.A.  
RC STRAIN=ATCC 23769;  
RX MEDLINE=94131945; PubMed=8300521;  
RA Standal R., Iversen T.-G., Coucheron D.H., Fjaervik E., Blatny J.M.,  
RT Valla S.;  
RT "A new gene required for cellulose production and a gene encoding  
RT cellulolytic activity in *Acetobacter xylinum* are colocalized with the  
RL bce operon.";  
RN J. Bacteriol. 176:665-672(1994).  
[5]  
RP FUNCTION.  
RC STRAIN=ATCC 53582;  
RX MEDLINE=90202941; PubMed=2138620;  
RA Lin F.C., Brown R.M. Jr., Drake R.R. Jr., Haley B.E.;  
RT "Identification of the uridine 5'-diphosphoglucose (UDP-Glc) binding  
RL subunit of cellulose synthase in *Acetobacter xylinum* using the  
RN photoaffinity probe 5-azido-UDP-Glc.";  
RP J. Biol. Chem. 265:4782-4784(1990).  
[6]  
RP 3D-STRUCTURE MODELING, AND MUTAGENESIS OF ASP-188; ASP-189; ASP-236;  
RC ASP-333; GLN-369; ARG-370 AND ARG-372.  
RX STRAIN=ATCC 23769;  
RA MEDLINE=21324818; PubMed=11430986;  
RT Saxena I.M., Brown R.M. Jr., Dandekar T.;  
RT "Structure-function characterization of cellulose synthase:  
RL relationship to other glycosyltransferases.";  
RN Phytochemistry 57:1135-1148(2001).  
[7]  
RP REVIEW ON DOMAIN ARCHITECTURE.  
RC MEDLINE=95189716; PubMed=7883697;  
RX Saxena I.M., Brown R.M. Jr., Fevre M., Geremia R.A., Henriessat B.;  
RA "Multidomain architecture of beta-glycosyl transferases: implications  
RT for mechanism of action.";  
RN J. Bacteriol. 177:1419-1424(1995).  
-1- FUNCTION: Bifunctional protein comprised of a catalytic subunit  
and a regulatory subunit. The catalytic subunit of cellulose  
synthase polymerizes uridine 5'-diphosphate glucose to cellulose  
in a processive way. The thick cellulosic mats generated by this  
enzyme probably provide a specialized protective environment to  
the bacterium. The regulatory subunit binds bis-(3'-5') cyclic  
dianglylic acid (C-di-GMP).  
-1- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP  
+ {(1,4)-beta-D-glucosyl}(N+1).  
-1- COFACTOR: Magnesium (By similarity).  
-1- ENZYME REGULATION: Activated by C-di-GMP.  
-1- PATHWAY: Bacterial cellulose biosynthesis.  
-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
(Potential).

```
CC CC -1- INDUCTION: Cellulose is produced at a linear rate with respect to
CC cell growth when O(2) is present.
CC -1- DOMAIN: There are two conserved domains in the globular part of
CC the catalytic subunit: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysts and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycane chain
CC in the active site.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC GLYCOSYLTRANSFERASE FAMILY 2.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
CC FAMILY.
CC -1- CAUTION: Was originally (Ref.1) thought to be two separate ORFs
CC named acsa and acsb, due to a frameshift in position 678.
CC -----
CC This swiss-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outpost on
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54676; CAA38487.1; ALT FRAME.
DR EMBL; X54676; CAA38488.1; ALT_FRAME.
DR EMBL; M96060; AAA16971.1; -.
DR InterPro; IPR003919; CellSynth_A.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF03170; BCSB; 1.
DR Pfam; PF00535; Glycos transf_2; 1.
DR PRINTS; PR01439; CELLSTHASEA.
KM Cellulose biosynthesis; Transferase; Glycosyltransferase;
Transmembrane; Inner membrane.
FT DOMAIN 1 741 CATALYTIC.
FT DOMAIN 742 1550 CYCLIC DI-GMP BINDING.
FT DOMAIN 147 240 CATALYTIC SUBDOMAIN A.
FT DOMAIN 317 377 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 423 443 POTENTIAL.
FT TRANSMEM 468 488 POTENTIAL.
FT TRANSMEM 507 527 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 1513 1533 POTENTIAL.
FT ACT_SITE 189 189 POTENTIAL.
FT ACT_SITE 333 333 POTENTIAL.
FT BINDING 236 236 SUBSTRATE (POTENTIAL).
FT BINDING 238 238 SUBSTRATE (POTENTIAL).
FT MUTAGEN 188 188 D->N: DECREASE IN ACTIVITY.
FT MUTAGEN 188 188 D->P: LOSS OF ACTIVITY.
FT MUTAGEN 189 189 D->Y: LOSS OF ACTIVITY.
FT MUTAGEN 236 236 D->Y: LOSS OF ACTIVITY.
FT MUTAGEN 333 333 D->R: LOSS OF ACTIVITY.
FT MUTAGEN 369 369 Q->M: LOSS OF ACTIVITY.
FT MUTAGEN 370 370 R->P: LOSS OF ACTIVITY.
FT MUTAGEN 370 370 R->Q: DECREASE IN ACTIVITY.
FT MUTAGEN 372 372 R->A: LOSS OF ACTIVITY.
SQ SEQUENCE 1550 AA; 168161 MW; 63AB8952BC39E961 CRC64;

Query Match 3.4%; Score 201; DB 1; Length 1550;
Best Local Similarity 19.8%; Pred. No. 1.3e-06;
Matches 156; Conservative 85; Mismatches 212; Indels 334; Gaps 37;
```



Db 151 VDI FVPTYN---EELSI VRLTVLSGIDWPPEKVRVHLLDDG----- 190  
QY 412 KWVFCRKRYNIEPRAPENWFQOKIDYLDKVAANFVRERRAMKREYEEFKVRINALVAKA 471  
Db 191 -----RRPEF-----AAFAAE----- 201  
QY 472 QKVPEEGWTMODGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPQYN 531  
Db 202 -----CGAN-----YIARPT---N 212  
QY 532 HHKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMM-DPLL GKVCYVQ 590  
Db 213 EHAKAGNLNYAIG---HTDGDYILIFDCDH-VPTRAFQLTMGMVEDP---KIALMQ 263  
QY 591 FPORFIDIDRHDYANRNVFEDINMKGLDGIQPIYVGTGCVFRQALYGYDAPKTKP 650  
Db 264 TPHHFYSPD-----PQORNL SAGY-----RTP 285  
QY 651 PSRTCNCWPKWCFCGCCFGNRKQKTKTKTEKKLLFKKEENQSPAYALGEIDEAABG 710  
Db 286 PE-----GN-----LFY----- 292  
QY 711 AENEKAGIVNOQKLEKFGQSVFVTSTLLENGT LKSA SPASL LKKA I HVISCGYEDKT 770  
Db 293 -----GVVD---GNDFWDATFF-----CGSCA I LRRTAIEQIG----- 323  
QY 771 DWGKEIGWYGSVTEDILGFKNHCHGWRSIYCI PKRVA FKGS-APLNLSDRLHQLRWMA 829  
Db 324 -----GFATQTVTEDAHTAL KQRLGWSTAYL---RIPLAGGLATERLILHIGQVRWMA 374  
QY 830 LGSIEIFFSNHCPLWYGGGKFLERFSYINSIVPWT SIPLAYCTLPALCILLTGKFI 889  
Db 375 RGM LQIFRIDN-PL---FGRGLSWGQRLCYLSAMTSFLFAVPRVIFLSSPLAFRFGQNI 430  
QY 890 ---TPELN NVASLWMSLFCIFATSILEMRMSGVG---IDWWRNEQFW-VIGVSSH 941  
Db 431 IASR-----LALLAYAI PHM-FHAVGTASKINKGWR-YSPWSEVYETTMA 474  
QY 942 LFAVFOGLLKVIAGVDTSPYTSKGGDDE---EFSELYTFKWTLLIPPTLLILNFI G 997  
Db 475 LFLVRYTVITLTLSPSRGKFNVTDKGLLEKGYFDLGAVY-----PN IILGLIMFGG 525  
QY 998 VVAGVSNAINNGYE-SWGPL-----FGKLFFAFWVIHLYPFLKGL-VGRQ-----NR 1043  
Db 526 LARGV-----YELSFGLDQIAERAYL LNSAWAMLSL I ILLA IAVGRETOQRNSHR 578  
QY 1044 TPTIVIV 1050  
Db 579 IPATIPV 585

RESULT 13  
BCA3\_ACEXY  
ID\_BCA3\_ACEXY STANDARD; PRT; 745 AA.  
AC Q9WX61;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cellulose synthase 1 catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSAI.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Acetobacteraceae; Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 7664 / IFO 13693;  
RX MEDLINE=99310341; PubMed=10382968;  
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,  
RA Inoue Y.;  
RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM  
RT 7664: implication of a novel set of cellulose synthase genes.";  
RL DNA Res. 6:109-115(1999).

CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose. The thick cellulosic  
CC mats generated by this enzyme probably provide a specialized  
CC protective environment to the bacterium (By similarity).  
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP  
CC + {(1,4)-beta-D-glucosyl}(N+1).  
CC -!- COFACTOR: Magnesium (By similarity).  
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
CC (c-di-GMP) (By similarity).  
CC -!- PATHWAY: Bacterial cellulose biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -!- DOMAIN: There are two conserved domains in the globular part of  
CC the catalytic subunit: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB015802; BAA77585.1; -.  
CC InterPro; IPR003919; Cellynth\_A.  
CC DR InterPro; IPR001173; Glycos\_transf\_2.  
CC DR Pfam; PF00535; Glycos\_transf\_2; 1.  
CC DR PRINTS; PR01439; CELLSNTHASEA.  
CC KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
CC Transmembrane; Inner membrane.  
CC KM  
CC FT DOMAIN 147 240 CATALYTIC SUBDOMAIN A.  
CC FT DOMAIN 317 377 CATALYTIC SUBDOMAIN B.  
CC FT TRANSMEM 29 49 POTENTIAL.  
CC FT TRANSMEM 106 126 POTENTIAL.  
CC FT TRANSMEM 153 173 POTENTIAL.  
CC FT TRANSMEM 407 427 POTENTIAL.  
CC FT TRANSMEM 430 450 POTENTIAL.  
CC FT TRANSMEM 468 488 POTENTIAL.  
CC FT TRANSMEM 515 535 POTENTIAL.  
CC FT TRANSMEM 547 567 POTENTIAL.  
CC FT TRANSMEM 649 669 POTENTIAL.  
CC FT ACT\_SITE 189 189 POTENTIAL.  
CC FT ACT\_SITE 333 333 POTENTIAL.  
CC FT BINDING 236 236 SUBSTRATE (POTENTIAL).  
CC FT BINDING 238 238 SUBSTRATE (POTENTIAL).  
CC SQ SEQUENCE 745 AA; 83518 MM; 57EA0457A226F815 CRC64;  
  
Query Match 3.3%; Score 194; DB 1; Length 745;  
Best Local Similarity 18.7%; Pred. No. 1.5e-06;  
Matches 156; Conservative 88; Mismatches 221; Indels 370; Gaps 36;  
  
QY 276 IIRLVVLGFFPHRYVMHPVNDAFALWLISVICEIMFAMSWILDQ---FPKWF----- 326  
Db 53 MIVALVCVAVPFIVGRKRSRRTQVLEVISALVSLRY-LTWRLTETIDFDTWQGI LGVT 111  
QY 327 -IERETY-LDRISLRFDEKQPSQLAP-----IDFFVSTVDP LKEPPLVTTNT 372  
Db 112 LLLAE L YALYMLFLSYQTISPLHRAPLPANPDWPTVDIFIPYD---EALSI VRLT 168  
QY 373 VLSILSVDPVDPKVSQVSDGAAMLTFEALSETSEFAKKWVPFCRKNIEPRAPENWFQ 432  
Db 169 VLGALGIDWPPDKVNYILDDG----- 190  
QY 433 QKIDYLDKVAANFVRERRAMKREYEEFKVRINALVAKAQKVPEEGWTMODGTPWPGNNV 492  
Db 191 -----RR-----EEFARFAEACGAR----- 205

```

QY 493 RDHPGMIQVFLGSGGLDCEGNELPLRVYVSREKRPQYNHKKAGANALVRVSAVLTTNA 552
Db 206 -----YIARPD---NAHAKAGINLYAIK---HTTG 229
QY 553 PYLLNLDCDHYNNSKAIKEAMCFMMDPLGKVCYVQFPQRFIDGRHRYANRVVFF 612
Db 230 DHILLDCDH-IPTRAFLOISMGMVVS---DSNIALLOTPHHFYSPD----- 272
QY 613 DINMGDLGIGPIYVGTGCVFRQALYGDAPKTKPPSRTCNCPKWCFCGCCFGNRK 672
Db 273 -----PQQRNLAVGY-----RTPPE-----GN-- 289
QY 673 OKKTKPKTEKKLLFFKKEENOSPAYALGEIDEAAPGAENEKAGIVNOCKLEKKGQSS 732
Db 290 -----LFY-----GVIGD---GNDPMDAT 305
QY 733 VEVVTSTLLENGGTLKSPASLKEAIVISCGYEDKTDWKEIGMTYGSVTEIDLTGFK 792
Db 306 FF-----CGSCAILRKAIEIG-----GFATETVTEDAHTALR 339
QY 793 MCHGWRSIYC-IPKRVAFKGSAPLNLSDRLHQLRWALGSEIFFSNHCPLWYGGGL 851
Db 340 MQRKGWSTAVLRIP---LASGLATERLITHTGQRMWRMARGMIQIFRVNPMI---GSGI 392
QY 852 KFLERFSYINSIVYPMWTSIPLAYCTLP-----AICLLTKRFTPELNNVASL 899
Db 393 KLGRLCYLSAMTSFFFAIPRVIFLASPLAFIFSQNIIAASPLAVGYAIPHM----- 446
QY 900 WFMSLFICIFATSILEMWMSGVGIDDMWRNDFWVIGVSSHLEAVFQGLKVIAGVD- 958
Db 447 -FHSI-----ATA-----AKVKNKGR-VSFW-----SEVETVVALFLVRVTIVTM 485
QY 959 -----SFTVTSKGG--DDEEFSELYTFKWTLLIPTTLLLNFIIVAVGSNAINGY 1010
Db 486 LFPSKKGKFNVTGKGVLEREEFDLTATY-----PNIIFAIIMALGLRG----- 529
QY 1011 ESWGPLFGKLFPAFWVIHLVPLKGLVGRQNRPTIVIVMS---ILLASIFSL 1061
Db 530 -----LYALIFQHLDI-----SERAYALNCIWSVISLIIIMAVISV 566

```

RESULT 14

ID	CBL_HUMAN	STANDARD;	PRT;	906 AA.
AC	P22681;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	CBL E3 ubiquitin protease (EC 6.3.2.-) (Signal transduction protein CBL) (Proto-oncogene c-CBL).			
GN	CBL OR CBL2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91232862; PubMed=2030914;			
RA	Blake T.J., Shapiro M., Morse H.C. III, Langdon W.Y.;			
RT	"The sequences of the human and mouse c-cbl proto-oncogenes show			
RT	v-cbl was generated by a large truncation encompassing a proline-rich			
RL	domain and a leucine zipper-like motif."			
RN	OncoGene 6:653-657(1991).			
RN	[2]			
RP	FUNCTION.			
RX	MEDLINE=99445925; PubMed=10514377;			
RA	Joazeiro C.A., Wing S.S., Huang H.-K., Levenson J.D., Hunter T.,			
RA	Liu Y.-C.;			
RT	"The tyrosine kinase negative regulator c-Cbl as a RING-type,			
RT	E2-dependent ubiquitin-protein ligase."			
RL	Science 286:309-312(1999).			
RN	[3]			

```

RP INTERACTION WITH SLA AND ZAP70, AND MUTAGENESIS OF GLY-306.
RX MEDLINE=99380595; PubMed=10449770;
RA Tang J., Sawadkowsky S., Chang J.-H., Burakoff S.J.;
RT "SLAP, a dimeric adapter protein, plays a functional role in T cell
RL receptor signaling.";
RN Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).
RP [4]
RP INTERACTION WITH SLA2.
RX MEDLINE=21553259; PubMed=11696592;
RA Holland S.J., Liao X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P.,
RA Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M.,
RA Yu S., Chan E., Wu X., Li C., Weisetschlager M., Aversa G.,
RA Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G.,
RA Mancebo H.S.Y., Wu J.;
RT "Functional cloning of Src-like adapter protein-2 (SLAP-2), a novel
RT inhibitor of antigen receptor signaling.";
RL J. Exp. Med. 194:1263-1276(2001).
RP [5]
RP INTERACTION WITH CD2AP.
RX MEDLINE=21265017; PubMed=11067845;
RA Kirsch K.H., Georgescu M.M., Shishido T., Langdon W.Y., Birge R.B.,
RA Hanafusa H.;
RT "The adapter type protein CBL/CD2AP binds to the proto-oncogenic
RT protein c-Cbl through a tyrosine phosphorylation-regulated Src
RL homology 3 domain interaction.";
RL J. Biol. Chem. 276:4957-4963(2001).
RP [6]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 47-350.
RX MEDLINE=99176421; PubMed=10078535;
RA Meng W., Sawadkowsky S., Burakoff S.J., Eck M.J.;
RT "Structure of the amino-terminal domain of Cbl complexed to its
RT binding site on ZAP-70 kinase.";
RL Nature 398:84-90(1999).
RP [7]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 47-434 IN COMPLEX WITH ZAP70
RP AND UBE2L3.
RX MEDLINE=20419298; PubMed=10966114;
RA Zheng N., Wang P., Jeffrey P.D., Pavletich N.P.;
RT "Structure of a c-Cbl-UbcH7 complex: RING domain function in
RT ubiquitin-protein ligases.";
RL Cell 102:533-539(2000).
CC -1- FUNCTION: Participates in signal transduction in hematopoietic
CC cells. Adapter protein that functions as a negative regulator of
CC many signaling pathways that start from receptors at the cell
CC surface. Acts as an E3 ubiquitin-protein ligase, which accepts
CC ubiquitin from specific E2 ubiquitin-conjugating enzymes, and then
CC transfers it to substrates promoting their degradation by the
CC proteasome. Recognizes activated receptor tyrosine kinases,
CC including PDGFR, EGF and CSF1, and terminates signaling.
CC -1- SUBUNIT: Associates with NCK via its SH3 domain. The
CC phosphorylated C-terminus interacts with CD2AP via its second SH3
CC domain. Binds to ZAP70 and UBE2L3. Interacts with adapters SLA and
CC SLA2.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DOMAIN: The RING-type zinc finger domain mediates binding to an E2
CC ubiquitin-conjugating enzyme.
CC -1- PTM: Phosphorylated on tyrosine residues.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 1 UBA domain.
CC -1- DATABASE: NAME=Atlas Genet. CytoGenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/CBLID171.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X57110; CAA40393.1; -
CC PIR; A43817; A43817.
CC PDB; 1B47; 27-APR-99.

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:00:29 ; Search time 45 Seconds

(without alignments)  
6227.670 Million cell updates/sec

Title: US-09-720-383C-10

Perfect score: 5918

Sequence: 1 MEASAGLVAGSHNRNRLVVI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5918	100.0	1086	10 Q9LLI3	Q91113 zea mays (m
2	5408.5	91.4	1059	10 Q9LLI4	Q91114 zea mays (m
3	5228	88.3	1094	10 Q9LLI2	Q91112 zea mays (m
4	4333.5	73.2	1091	10 Q93XQ1	Q93XQ1 nicotiana a
5	4278.5	72.3	1084	10 Q48947	Q48947 arabidopsis
6	4262	72.0	1084	10 Q9FGF9	Q9fgf9 arabidopsis
7	4209.5	71.1	1088	10 Q9SJ22	Q9sj22 arabidopsis
8	4204.5	71.0	1069	10 Q9FIB9	Q9fib9 arabidopsis
9	4202	71.0	1081	10 Q65338	Q65338 arabidopsis
10	3984	67.3	1075	10 Q9LLI9	Q91119 zea mays (m
11	3977.5	67.2	1074	10 Q9LLI8	Q91118 zea mays (m
12	3967.5	67.0	1026	10 Q9SWM6	Q9swm6 arabidopsis
13	3963.5	67.0	1026	10 Q9XHP6	Q9xhp6 arabidopsis
14	3955.5	66.8	1032	10 Q8GSW2	Q8gsW2 populus tre
15	3945.5	66.7	1081	10 Q48946	Q48946 arabidopsis
16	3929.5	66.4	1067	10 Q9XGX6	Q9xgx6 gossypium h

17	3911	66.1	1076	10 Q9LLI5	Q91115 zea mays (m
18	3905.5	66.0	1077	10 Q9LLI6	Q91116 zea mays (m
19	3898.5	65.9	1065	10 Q9FHK6	Q9fhk6 arabidopsis
20	3886	65.7	1042	10 Q81649	Q81649 populus x c
21	3880.5	65.6	1079	10 Q9LLI1	Q91111 zea mays (m
22	3875.5	65.5	1065	10 Q48948	Q48948 arabidopsis
23	3835.5	64.8	1063	10 Q9AV71	Q9av71 oryza sativ
24	3826	64.7	1129	10 Q8LK26	Q8lk26 mesoetanium
25	3810	64.4	1055	10 Q8GZN8	Q8gzN8 arabidopsis
26	3804.5	64.3	1043	10 Q9FNC3	Q9fnc3 arabidopsis
27	3782	63.9	1065	10 Q9SKJ5	Q9skj5 arabidopsis
28	3476.5	58.7	978	10 Q9AXK0	Q9axK0 zinnia eleg
29	3469.5	58.6	974	10 P93155	P93155 gossypium h
30	3421	57.8	985	10 Q9C528	Q9c5z8 arabidopsis
31	3420	57.8	985	10 Q8LPK5	Q8lpk5 arabidopsis
32	3410.5	57.6	974	10 Q8MW0	Q8mw0 gossypium h
33	3332.5	56.3	978	10 Q81368	Q81368 populus tre
34	3329.5	56.3	958	10 Q9SN37	Q9sn37 arabidopsis
35	3318	56.1	821	10 Q9LLI7	Q91117 zea mays (m
36	3165	53.5	939	10 Q943H3	Q943h3 oryza sativ
37	2907	49.1	685	10 P93156	P93156 gossypium h
38	2330.5	39.4	1181	10 Q9SRW9	Q9srW9 arabidopsis
39	2326	39.3	1145	10 Q9M9M4	Q9m9M4 arabidopsis
40	2324.5	39.3	1104	10 Q8GJZ9	Q8guZ9 populus tre
41	2321.5	39.2	1170	10 Q9LHZ7	Q9lhZ7 oryza sativ
42	2311	39.1	1145	10 Q9LFLO	Q91f10 arabidopsis
43	2293.5	38.8	1127	10 Q8W3F9	Q8w3f9 oryza sativ
44	2256	38.1	1111	10 Q9SZL9	Q9sz19 arabidopsis
45	2253	38.1	1127	10 Q93XQ0	Q93XQ0 nicotiana a

ALIGNMENTS

RESULT 1	ID	Q9LLI3	PRELIMINARY;	PRT; 1086 AA.
AC	Q9LLI3;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DB	Cellulose synthase-7.			
GN	CESA-7.			
OS	Zea mays (Maize).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	PACCAD clade; Panicoideae; Andropogoneae; Zea.			
OX	NCBI_Taxid=4577;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20398328; PubMed=10938350;			
RA	Holland N., Holland D., Helentjaris T., Dhugga K.S.,			
RA	Xoconostle-Cazares B., Delmer D.P.;			
RT	"A comparative analysis of the plant cellulose synthase (Cesa) gene family.";			
RL	Plant Physiol. 123:1313-1324(2000).			
DR	EMBL; AF200531; AAF89967.1; -.			
DR	InterPro; IPR005150; Cellulose_synt.			
DR	InterPro; IPR001841; Znf_ring.			
DR	Pfam; PF03552; Cellulose_synt; 1.			
DR	SMART; SM00184; RING; 1.			
DR	PROSITE; PS50089; ZF_RING 2; 1.			
DR	SEQUENCE 1086 AA; 122608 MW; DD03C73ABD13E2EB CRC64;			

Query Match 100.0%; Score 5918; DB 10; Length 1086;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 MEASAGLVAGSHNRNRLVIRDDPGPKPREQNGQVCQICGDDVGLAPGDPFVACNE 60
Db	1 MEASAGLVAGSHNRNRLVIRDDPGPKPREQNGQVCQICGDDVGLAPGDPFVACNE 60
Qy	61 CAFPVCRDCEYERREGTQNCQCKTRYKRLKGQCVRTGDEBEDGVDDLDFNFWNDGHS 120

Db	61	CAFPVCRDCYERREGTQNCPCQCKTRYKRLKGQCRVYTGDEEEDGVDDLNEFNWDGHS	120
Qy	121	QSVAESMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQWDDIPPEQHALVSEFMGG	180
Db	121	QSVAESMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQWDDIPPEQHALVSEFMGG	180
Qy	181	GKRIHPLPYADPSLPVQPRSMDBSKDLAAYGYSVANKERMENMKQROERMHOTGDDGG	240
Db	181	GKRIHPLPYADPSLPVQPRSMDBSKDLAAYGYSVANKERMENMKQROERMHOTGDDGG	240
Qy	241	DDGDDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFHRYRVMPVNDAPA	300
Db	241	DDGDDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFHRYRVMPVNDAPA	300
Qy	301	LWLISVCEIWFAMSWILDQFPKFPPIERETYLDRLSLRFDEKQPSQAPIDFVSTVD	360
Db	301	LWLISVCEIWFAMSWILDQFPKFPPIERETYLDRLSLRFDEKQPSQAPIDFVSTVD	360
Qy	361	PLKEPPLVTNTVLSTLSVDYPVDKVSQVSDGAAMLTFEALSETSEFAKKWVPECKRY	420
Db	361	PLKEPPLVTNTVLSTLSVDYPVDKVSQVSDGAAMLTFEALSETSEFAKKWVPECKRY	420
Qy	421	NIEBRABEWYFOQKIDYLKQKVAANFVRERBRAMKREYEEFKVRINALVAKAQKVPBGWT	480
Db	421	NIEBRABEWYFOQKIDYLKQKVAANFVRERBRAMKREYEEFKVRINALVAKAQKVPBGWT	480
Qy	481	MODGTWPNGNNVRDHPGMIQVFLQSGGLDCEGNELPRLVYVSREKRPNGYHHKKAGAMN	540
Db	481	MODGTWPNGNNVRDHPGMIQVFLQSGGLDCEGNELPRLVYVSREKRPNGYHHKKAGAMN	540
Qy	541	ALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFQIDR	600
Db	541	ALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFQIDR	600
Qy	601	HDRYANRNWVFEDINMKGLDGIQGPITYVGTGCVFRQALYGDAPKTKPPSRTCNCWPK	660
Db	601	HDRYANRNWVFEDINMKGLDGIQGPITYVGTGCVFRQALYGDAPKTKPPSRTCNCWPK	660
Qy	661	WCFCCCFGNRKOKTKTKPKTEKKLLFKKEENOSPAYALGEIDEAPGAENEKAGIVN	720
Db	661	WCFCCCFGNRKOKTKTKPKTEKKLLFKKEENOSPAYALGEIDEAPGAENEKAGIVN	720
Qy	721	QOKLEKFGQSSVFVTSTLLENGSTLKSASPASLKEAIVHISCGYEDKTDMGKEIGWY	780
Db	721	QOKLEKFGQSSVFVTSTLLENGSTLKSASPASLKEAIVHISCGYEDKTDMGKEIGWY	780
Qy	781	GSVTEIDILTFKMHCHGWSIYCI PKRVAFKGSAPLNSDRLHOVLRWALGSIEIFPSNH	840
Db	781	GSVTEIDILTFKMHCHGWSIYCI PKRVAFKGSAPLNSDRLHOVLRWALGSIEIFPSNH	840
Qy	841	CPLWYGGGLKFLERFSYINSIYVWTSIPLAYCTLPALCLLTGKFITPELNNVASLW	900
Db	841	CPLWYGGGLKFLERFSYINSIYVWTSIPLAYCTLPALCLLTGKFITPELNNVASLW	900
Qy	901	FMSLFCIFATSILEMWMSGVGDIDMWRNEQFWVIGVSSHFAVQGLLKVIAGVTSF	960
Db	901	FMSLFCIFATSILEMWMSGVGDIDMWRNEQFWVIGVSSHFAVQGLLKVIAGVTSF	960
Qy	961	TVTSKGDDDEESELTYTFKWTLLIPPTLLLLNFIGVAVGSNAINNGYESWGPLFGKL	1020
Db	961	TVTSKGDDDEESELTYTFKWTLLIPPTLLLLNFIGVAVGSNAINNGYESWGPLFGKL	1020
Qy	1021	FFAFWVIVHLYPFLKGLVGRQNRPTIIVVWSILASIFSLMWRIDPFLAKDDGPLLEE	1080
Db	1021	FFAFWVIVHLYPFLKGLVGRQNRPTIIVVWSILASIFSLMWRIDPFLAKDDGPLLEE	1080
Qy	1081	CGLDCN 1086	
Db	1081	CGLDCN 1086	

RESULT 2

Q9LL14	PRELIMINARY;	PRT;	1059 AA.
ID Q9LL14			
AC Q9LL14;			
DT 01-OCT-2000 (Tremblrel. 15, Created)			
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)			
DE Cellulose synthase-6.			
GN CESA-6.			
OS Zea mays (Maize).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.			
OX NCBI_TaxID=4577;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=20398328; PubMed=10938350;			
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,			
RA Xocconostle-Cazares B., Delmer D.F.;			
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene family."			
RL Plant Physiol. 123:1313-1324 (2000).			
DR EMBL; AF200530; AAF89966.1; -.			
DR InterPro; IPR005150; Cellulose_synt.			
DR InterPro; IPR001841; Znf ring.			
DR Pfam; PF03552; Cellulose_synt; 1.			
DR SMART; SM00184; RING; 1.			
DR PROSITE; PS50089; ZF_RING_2; 1.			
SQ SEQUENCE 1059 AA; 120044 MW; 289DA26B25232249 CRC64;			

Query Match 91.4%; Score 5408.5; DB 10; Length 1059;  
Best Local Similarity 94.0%; Pred. No. 0;  
Matches 995; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

Qy	33	EONGVQCICGDDVGLAPGDDPFAVCNECAFVPCRDCEYERREGTQNCPCQCKTRYKRLK	92
Db	3	QNRGQVCQICGDDVGRNPDGEPFVACNECAFPICRDCYERREGTQNCPCQCKTRFKRFK	62
Qy	93	GCQRTVGDEEBDGVDDLNEFNW-DGHDSQSVAESMLYGHMSYGRGDPNGAPQAFQLN	151
Db	63	GCAVPGDEEBDGVDDLNEFNWSDKHDSQVLAESMLHAHMSYGRGADLDGVQFHPHP	122
Qy	152	NVPLLTNGQWDDIPPEQHALVPSFMGGGKRIHPLPYADPSLPVQPRSMDBSKDLAAYG	211
Db	123	NVPLLTNGQWDDIPRDQHALVPSFVGCGGKRIHPLPYADPNLPVQPRSMDBSKDLAAYG	182
Qy	212	YGSVANKERMENMKQROERMHOTGNDGGDGDADLPLMDEARQQLSRKIPLPSSQINP	271
Db	183	YGSVANKERMENMKQROERMHOTRNDGGDGDADLPLMDEARQQLSRKIPLPSSQINP	242
Qy	272	YRMIIIRLVLVGFFHYRVMPVNDAPALMLISVCEIWFAMSWILDQFPKFPPIERET	331
Db	243	YRMIIIRLVLVCFHFHYRVMPVNDAPALMLISVCEIWFAMSWILDQFPKFPPIERET	302
Qy	332	YLDRLSLRFDEKQPSQAPIDFVSTVDBLKEPPLVTNTVLSTLSVDYPVDKVSQVVS	391
Db	303	YLDRLSLRFDEKQPSQAPVDFVSTVDBLKEPPLVTANTVLSTLSVDYPVDKVSQVVS	362
Qy	392	DDGAAMLTFEALSETSEFAKKWVPFCRKYNIEBRABEWYFOQKIDYLKQKVAANFVRER	451
Db	363	DDGAAMLTFEALSETSEFAKKWVPFCRKYNIEBRABEWYFOQKIDYLKQKVAANFVRER	422
Qy	452	AMKREYEEFKVRINALVAKAQKVPBEGWTMODGTWPNGNNVRDHPGMIQVFLQSGGHDV	511
Db	423	AMKREYEEFKVRINALVAKAQKVPBEGWTMODGTWPNGNNVRDHPGMIQVFLQSGGHDV	482
Qy	512	EGNELPRLVYVSREKRPNGYHHKKAGANALVRSVAVLTNAPYLLNLDCHYINNSKAIK	571
Db	483	EGNELPRLVYVSREKRPNGYHHKKAGANALVRSVAVLTNAPYLLNLDCHYINNSKAIK	542
Qy	572	EAMCFMMDPLGKKVCYVQFPQRFQIDRDRYANRNWVFEDINMKGLDGIQGPITYVGTG	631
Db	543	EAMCFMMDPLGKKVCYVQFPQRFQIDRDRYANRNWVFEDINMKGLDGIQGPITYVGTG	602



QY 632 CVFRQALYGDAPKTKPPSRTCNWPKWCFCGCCFNGNRKÖK--KTTYPKTEKKLLF 688  
Db 603 CVFRQALYGDAPKTKPPSRTCNWPKWCICCCCFNGNRKTKKTKTSKPKFKIKKL- 661  
QY 689 FKKEENQSPAYALGEIDEAPGAENKAGIVNQÖKLEKKFGÖSSVFVTSTLLENGGTLKS 748  
Db 662 FKKEENQAPAYALGEIDEAPGAENKASIVNQÖKLEKKFGÖSSVFVASTLLENGGTLKS 721  
QY 749 ASPASLKEAIVHISCGYEDKTDWKEIGWYGSVTEDILTGFKMHCHGWSIYCIPIKRV 808  
Db 722 ASPASLKEAIVHISCGYEDKTGWGKDIGWYGSVTEDILTGFKMHCHGWSIYCIPIKRA 781  
QY 809 AFKGSAPLNLSDRLHÖVLRWALGSIEIFFSNHCPLMWYGGLKFLERFSYINSIVYPWT 868  
Db 782 AFKGSAPLNLSDRFHÖVLRWALGSIEILFSNHCPLMWYGGLKFLERFSYINSIVYPWT 841  
QY 869 SIPLAYCTLPALICLLTGKFITPELNNVASLWMSLFICIFATSILEMRWSGVGIDDWR 928  
Db 842 SIPLAYCTLPALICLLTGKFITPELNNVASLWMSLFICIFATSILEMRWSGVGIDDWR 901  
QY 929 NEQFWIVGVSSHFAVÖGLKVIAGVDTSFVTYSKGDDSEFSSELYTFKWTLLIPT 988  
Db 902 NEQFWIVGVSSHFAVÖGLKVIAGVDTSFVTYSKGDDSEFSSELYTFKWTLLIPT 961  
QY 989 TLLLNFIGVAVGSNAINNGYESWGPLFGKLEFAFWIVHLYPFLKGLVGRONRTPTIV 1048  
Db 962 TLLLNFIGVAVGSNAINNGYESWGPLFGKLEFAFWIVHLYPFLKGLVGRONRTPTIV 1021  
QY 1049 IVWSILLASIFSLMWRIDPFLAKDDGPLLEEGLDCN 1086  
Db 1022 IVWSILLASIFSLMWRIDPFLAKDDGPLLEEGLDCN 1059

RESULT 3

Q9LLI2 PRELIMINARY; PRT; 1094 AA.  
AC Q9LLI2;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, last annotation update)  
DE Cellulose synthase-8.  
GN CESA-8.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20398328; PubMed=10938350;  
RA Holand N., Holand D., Helentjaris T., Dhugga K.S.,  
RA Xocoostle-Cazares B., Delmer D.P.;  
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene  
RT family."  
RL Plant Physiol. 123:1313-1324(2000).  
DR EMBL; AF200532; AAF89968.1; -;  
DR InterPro; IPR005150; Cellulose\_synt.  
DR InterPro; IPR001841; Znf ring.  
DR Pfam; PF03552; Cellulose\_synt; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
SQ SEQUENCE 1094 AA; 122575 MM; DFBBC18F49D23F5E CRC64;

Query Match 88.3%; Score 5228; DB 10; Length 1094;  
Best local Similarity 88.7%; Pred. No. 0;  
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

QY 1 MEASAGLVAGSHRNELVIRRD---GDPGPKPREQÖNQVCÖICGDDVGLAPGDPFVA 57  
Db 1 MEASAGLVAGSHRNELVIRRDRESGAAGGGAARRAEP-CÖICGDEVGVGFDGEPFVA 59  
QY 58 CNECAFVPCRDYCYERREGTÖNCPÖCKTRYKRLKGCÖQRTGDEEDGVDDLNEFN-WD 116

Db 60 CNECAFVPCRDYCYERREGSQACPÖCKTRYKRLKGCPRVAGDEEEDGVDDLBEFGLOD 119  
QY 117 G---HDSQVAESMLYGHMSYGRGDPNGAPÖAFÖLNPVPLLTNGÖMVDIIPPEÖHAL 172  
Db 120 GAHEDDPÖYVAESMLRAÖMSYGRGD---AHGFSVPVPNPVPLLTNGÖMVDIIPPEÖHAL 176  
QY 173 VPSFM---GGGKRHPPLPYADPSLPVÖPRSMDPKDLAAYGVGSVAWKERMENWKÖRÖE 229  
Db 177 VPSYMSGGGGGKRHPPLPYADPSLPVÖPRSMDPKDLAAYGVGSVAWKERMENWKÖRÖE 236  
QY 230 RMHÖTGNDGGGD-DGDDADPLMDÖARÖLSRKIPLPSSÖQNPYRMIIIRLVVLGFFFH 288  
Db 237 RLÖHVRSEGGGDWDGDDADPLMDÖARÖPLSRKVPISSSRINPYRMIIIRLVVLGFFFH 296  
QY 289 YRVMPVNDAPALWLISVCEIWPAMSILDÖPKWFPPIERETYLDRLSLPDKEGÖPSÖ 348  
Db 297 YRVMPVNDAPALWLISVCEIWPAMSILDÖPKWLPPIERETYLDRLSLPDKEGÖPSÖ 356  
QY 349 LAPIDFVSTVDPLKEPPLVTNTVLSLSVDYPVDKVSQVSDGAAMLTFEALSETSE 408  
Db 357 LAPIDFVSTVDPTKEPPLVTANTVLSLSVDYPVEKVSQVSDGAAMLTFEALSETSE 416  
QY 409 FAKKWVPCKRYNIEBRAPÖWYFÖKIDYLDKVAANFVRERAMKREYEBFKVRINALV 468  
Db 417 FAKKWVPCKRYNIEBRAPÖWYFÖKIDYLDKVAASFVRERAMKREYEBFKVRINALV 476  
QY 469 AKAÖKVPBEÖWTMÖDGTWPBGNNVRDHPGMIÖVPLGÖSGGLDCEGNEPLRLVYVSREKRP 528  
Db 477 AKAÖKVPBEÖWTMÖDGSWPBGNNVRDHPGMIÖVPLGÖSGGRDVEGNEPLRLVYVSREKRP 536  
QY 529 GYNHKKAGAMNALVRVSAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKVCY 588  
Db 537 GYNHKKAGAMNALVRVSAVLSNAYLLNLDCDHYINNSKAIKEAMCFMMDPLVGKVCY 596  
QY 589 VÖFPÖRFDGIDRHRDYNRNVPFDINMGLDGIÖGPIYVGTGCVFRQALYGDAPKTX 648  
Db 597 VÖFPÖRFDGIDKDRYANRNVPFDINMGLDGIÖGPIYVGTGCVFRQALYGDAPKTX 656  
QY 649 KPSRTCNWPKWCFCGCCFNGNRKÖKTKTPKTEKKLLFPKKEENQSPAYALGEIDBA 708  
Db 657 KPSRTCNWPKWCLSCCFSRNKÖKTKTPKTEKKRLFFKAENPSPAYALGEIDEGA 716  
QY 709 PGAENKAGIVNQÖKLEKKFGÖSSVFVTSTLLENGGTLKSASPASLKEAIVHISCGYED 768  
Db 717 PGADIEKAGIVNQÖKLEKKFGÖSSVFVASTLLENGGTLKSASPASLKEAIVHISCGYED 776  
QY 769 KTDWKEIGWYGSVTEDILTGFKMHCHGWSIYCIPIKRVAFKGSAPLNLSDRLHÖVLRW 828  
Db 777 KTDWKEIGWYGSVTEDILTGFKMHCHGWSIYCIPIKRVAFKGSAPLNLSDRLHÖVLRW 836  
QY 829 ALGSIEIFFSNHCPLMWYGGLKFLERFSYINSIVYPWTSIPLAYCTLPALICLLTGKF 888  
Db 837 ALGSIEIFFSKHCPLMWYGGLKFLERFSYINSIVYPWTSIPLAYCTLPALICLLTGKF 896  
QY 889 ITPELNNVASLWMSLFICIFATSILEMRWSGVGIDDWRNEQFWIVGVSSHFAVÖG 948  
Db 897 ITPELNNVASIWMALFICISVTGILEMRWSGVAIDDMWRNEQFWIVGVSAHLFAVÖG 956  
QY 949 LKVIAGVDTSFVTYSKGDDSEFSSELYTFKWTLLIPTLLLNFIGVAVGSNAINN 1008  
Db 957 LKVPAGIDTSFTVTSKAGDDEFSSELYTFKWTLLIPTLLLNFIGVAGISNAINN 1016  
QY 1009 GYESWGPLFGKLEFAFWIVHLYPFLKGLVGRONRTPTIVIVWSILLASIFSLMWRIDP 1068  
Db 1017 GYESWGPLFGKLEFAFWIVHLYPFLKGLVGRONRTPTIVIVWSILLASIFSLMWRIDP 1076  
QY 1069 FLAKDDGPLLEEGLDCN 1086  
Db 1077 FLAKSNGPLLEEGLDCN 1094

RESULT 4  
Q93XQ1

```
ID 093X01 PRELIMINARY; PRT; 1091 AA.
AC 093X01;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN CESA1.
CS Nicotiana alata (winged tobacco) (Persian tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4087;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Distil;
RX MEDLINE=21196092; Pubmed=11299383;
RA Doblin M.S., De Melis L., Newbiglin E., Bacic A., Read S.M.;
RT "Pollen Tubes of Nicotiana alata Express Two Genes from Different
RT beta-Glucan Synthase Families.";
RL Plant Physiol. 125:2040-2052(2001).
DR EMBL; AF304374; AAK49454.1; -.
DR InterPro; IPR005150; Cellulose_synt.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1091 AA; 122682 MW; 3FB8DF94B19D519B CRC64;

Query Match 73.2%; Score 4333.5; DB 10; Length 1091;
Best Local Similarity 71.9%; Pred. No. 0;
Matches 794; Conservative 123; Mismatches 153; Indels 35; Gaps 12;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCQICGDDVGLAPGDPFVACNE 60
Db 1 MDTKGRVLVAGSHNRNEFYINADVGRTSVKELSGQICQICGEIEVTVDGERPIACNE 60

QY 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGDEEDGVDDLDFNEFWGDHD- 119
Db 61 CAFPVCRDCYEYERREGNACPOCKTRFKRIKGSFRVDGDEDDDEFDDLDFHEFYHGNPR 120

QY 120 --SQSVAESMLYGHMSYRGCDPNGA-----PQAFQLNPVPLLTNGQWVDIPPE 168
Db 121 YMSEAAFSRL-----GRGTNNASGLTTPSEVDPA--LNSEIPLLTYGQEDDTISAD 172

QY 169 QHAL-VPSFMGGGKRIHPLPYADPSLPVQPRSMDBSKDLAAYGYGSVAMKEMENWKOR 227
Db 173 KHALIIPFM-GRGKVVHPVYSD-SMSLPRPRMDPKDLAVYGYGTVAWKERMEDWKKK 230

QY 228 QERMHQT---GNDGGGDDG---DDADLPLMDEARQQLSRKPLPSSQINPYRMIIIRL 280
Db 231 QNDKLQVVKHGGGKGGNDGDELDDPLPKMDEGRQPLSRKLPISSSRLSPYRLLIVRL 290

QY 281 VVLGFEEHYRVMHPVNDARFALWLISVCEIWFAMSWILDQPKMPIRETYLDRLSLRF 340
Db 291 AVVGLFFHYRITHPVNDAYALWLISICEIWFAMSWIFDQPKMFIYRETYLDRLSLRY 350

QY 341 DKEGQPSQLAPIDFFVSTVDPLKEPRLVTNTVLSLSDYPRVDKVSQVSDGAAMLTF 400
Db 351 EKEGKPSGLAPIDIFVSTVDPLKEPRLITANTVLSILAVDYPEDKVSQVSDGAAMLTF 410

QY 401 EALSETSEFAKKWVPFCCKRYNIEPRAPPEWYFOQKIDYLKDKVAANFVRERRAMGEYEEF 460
Db 411 EALSETSEFAKKWVPFCCKRYNIEPRAPPEWYFQKVDYLNKVKVPSFVRERRAMKDYEEF 470

QY 461 KVRINALVAKAQKVPBEQTMQDGTWPGNNVRDHPGMIQVFLGQSGGLDCEGNEPLRLV 520
Db 471 KVRINGLVATAQKVPEDGTMQDGTWPGNLVRDHPGMIQVFLGNDGVRDIEGNVLPRLI 530

QY 521 YVSREKRPGYNHKKAGANVALVRSVAVLTNAPYLLMLDCDHYINNSKAIKEAMCFMMDP 580
Db 531 YVSREKRPGFDDHKKAGANVALMRVSAVISNAPYLLNVDCHYINNSKALREAMCFMMDP 590

QY 581 LUGKVCYVQFPQRFDGIDRHDYRANRNVVFFDINMKGLDGIQGPITYVGTGCVERRQALY 640
```

```
Db 591 TSGKKICVQFPQRFDGIDRHDYRANRNVVFFDINMKGLDGIQGPITYVGTGCVERRQALY 650

QY 641 GYDAPKTKPPSRTCNCWPKWCFCCCFGNRKOKTKTPKTEKKLLFPKKEENQSPAYA 700
Db 651 GYDAPKTKPPSKTCNCWPKW-CCCCFGSRKKHKKKTKTKDNKKKT--KTKBASPIQIHA 706

QY 701 LGEIDBAAPGAENKAGIVNQOKLEKKFGQSSVFVTSTLLENGTLLKSASPASLLKEAII 760
Db 707 LENIEBEGIEGIDSEKATLMPQIKLEKFKGQSPVFVASTLLEDGIPPGATSASLLKEAII 766

QY 761 VISCGYEDKTDWCKEIGWIYGSVTEDILTGFKMCHGWRSTYCTIPKRVAFKGSAPLNSD 820
Db 767 VISCGYEDKTEWGREVGWIYGSVTEDILTGFKMCHGWRSVYCMFKRAPFKGSAPINLSD 826

QY 821 RLHQVLRNALGSIEIFFSNHCPWYGYGGLKFLERFSYINSIYVWTSIPLAYCTLPA 880
Db 827 RLHQVLRNALGSVEILLSKHCPWYGYGGLKFLERFSYINSVYPLTSLPLIAYCALPA 886

QY 881 ICLLTGKFTPELNNVASLWFMSLFCIFATSILEMWMSGVIGIDWWRNEQFVIGVSS 940
Db 887 VCLLTGKFTPEISNYASILFMGLFIMIAATSVLEMQMGVITIDWWRNEQFVIGVSS 946

QY 941 HLFAPFQGLKYIAGVDTSTFTVTSKGGDDEEFSELYTFKWTTLIPPTLLINFIGVVA 1000
Db 947 HLFAPFQGLKYLAGVSTSTFTVTSKADDEEFSELYTFKWTSLIPPTLLINIGVIV 1006

QY 1001 GVSNAINNGYSWGPIFGKLFAPFVIVHLYPFLKGLVGRONRPTIVWSILLASIFS 1060
Db 1007 GISDAINNGYDSWGPIFGRLFFALMWIVHLYPFLKGVMGROKVPITIVWSILLASIFS 1066

QY 1061 LLMWRIDPFLAKDDGPLLEBEGGLDC 1085
Db 1067 LLMWRVNPFTAR-GGLVLEVCGLDC 1090

RESULT 5
ID 048947 PRELIMINARY; PRT; 1084 AA.
AC 048947;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN H-A OR T22F8.250 OR AT4G39350.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=98111412; Pubmed=9445479;
RA Arioli T., Peng L., Beltzner A.S., Burn J., Wittke W., Herth W.,
RA Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,
RA Redmond J., Williams R.E.;
RT "Molecular analysis of cellulose biosynthesis in Arabidopsis.";
RL Science 279:717-720(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Bevan M., Zimmermann W., Grueneisen A., Wambutt R., Bancroft I.,
RA Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Zimmermann W., Grueneisen A., Wambutt R., Kalicki J., Wohldmann P.,
RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
```





Oy		294	PVNDALFALWLISVICIEMFAMSWILDOFPKMFPIERETYLDRLSLRFDKEGQPSQLAPID	353		
Db		297	PVKDAYALWLVISVICEIMFAVSFWLDQFPKWYPIERETYLDRLSLRYEKEGKPSGLSPVD	356		
Oy		354	FVVSTVDPLKEPPLVTNTVLISLSVDYPVDKVSCVSDDGAAMLTFEALSETSEFAKKW	413		
Db		357	VFVSTVDPLKEPPLTANTVLSILAVDYPVDKVACVSDDGAAMLTFEALSETAEARKW	416		
Oy		414	VPFCRKNIEPRAPENWFQOKIDYLLKDVAANFVRERRAMKREYEFPKRINALVAKAQK	473		
Db		417	VPFCCKYCIEPRAPENWFCHKMDYLKKNKHVPAFVRERRAMKRDYEEFKVINALTVAQAOK	476		
Oy		474	VPBEGWTMODGTPWPGNNVRDPHGMIOVFLLGOSGGILDCEGNELPRLVTVSREKRPNGNH	533		
Db		477	VPEDGMTODGTPWPBGNSVYRDHPGMIQVFLGSDGVADVENNELLRLVTVSREKRPFGDH	536		
Oy		534	KKAGAMNALVRVSAVLTNADPYLLNLDCDHYNNSKAIEAMCFMMDPLLGRKVCYQFPQ	593		
Db		537	KKAGAMNSLRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPOSGKKICYQFPQ	596		
Oy		594	RFDGIDRHDRYANRVNVEFDINMKGLDGIQGPITYGTGCVFRRQALYGDAPKTKKPSR	653		
Db		597	RFDGIDRHDRYSNRNVFFDINMKGLDGLQGPIYVGTCVFRQALYGFDAPKKKGBRK	656		
Oy		654	TNCNMPKMWCFCGCCFGRNKOKTTKPKTEKKULLFFKKEENSOPAYALGEIDE--AAPGA	711		
Db		657	TNCNMPKMCL--LCFGRNRKRAKTVAADKK---KNREASKOIHALENIEEGRTKGS	710		
Oy		712	ENEKAGIVNOQKLEKKFFGOSSVFTVSTLLENGCTLKSASPASLLKEAIIHVISCGYEDTD	771		
Db		711	NVEOSTEAMQMKEKKFFGQBPVVASARMENGAMARNASACILLKEAIQVISCGYEDTE	770		
Oy		772	WGKEIGWIYGSVEDILTGFPMCHGWRSIYCIIPKRVAFKGSAPLNLSDRLHQVLRNALG	831		
Db		771	WGKEIGWIYGSVEDILTGFPMHSHGWRSVYCTPKLAEFKGSAPINLSDRLHQVLRNALG	830		
Oy		832	SIEIFFSNMCPLMYGGGGLKFLERSYINSIYVPWTSIPLLAYCTLPAILCLTGKFIPTP	891		
Db		831	SVEIFLSRHCPIMYGGGGLKMLERLSYINSVVPWTSIPLIYCSLPAICLLTGKFIVP	890		
Oy		892	ELNNVASLWMFMSLFICIPATSILEMWSGVGIDDMMWRNEQFVWIGVSSHFAVFOGLIK	951		
Db		891	EISNYASILFMALFSSIAITGLEMQMGKVGIDDMMWRNEQFVWIGVSAHLFALFOGLIK	950		
Oy		952	VIAGVDTSFVTSSKGDEEFSELTYFKMTTLILPPTTLILNFIVGAVSNAINNGYE	1011		
Db		951	VLAGVDTNFTVTSKADDGEFSDLYLFKMTSLLIPMTLLINIVGIVGVSDAISNGYD	1010		
Oy		1012	SWGPLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPTIVIVMSILLASIFSLLWVRIDPFILA	10711		
Db		1011	SWGPLFGRLFFALMWIILHLYPLKGLLGKQDRMPTIIVVMSILLASITLLWVRVNPFIYA	1070C		
Oy		1072	KDDGPLLBECGLDC	1085		
Db		1071	K-GGPFILEICGLDC	1083		
RESULT 7						
ID	Q9SJ22	PRELIMINARY;	PRT; 1088 AA.			
AC	Q9SJ22;					
DT	01-MAY-2000	(TREMBLrel. 13, Created)				
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)				
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)				
DE	Putative cellulose synthase catalytic subunit.					
GN	AT2G21770.					
OS	Arabidopsis thaliana (Mouse-ear cress).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;					
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.					
NCBI_TaxId=3702;						
RN	[1]					

RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhagen G.P., Preuss D., Niernman W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant *Arabidopsis*  
RT *thaliana*.";   
RL Nature 402:761-768 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007019; AAD20396.1; -.  
DR InterPro; IPR005150; Cellulose\_synt.  
DR InterPro; IPR001841; znf\_ring.  
DR pfam; PF03552; Cellulose\_synt; 1.  
DR SMART; SMO0184; RING; 1.  
SQ SEQUENCE 1088 AA; 123446 MW; 060571118600DC9F CRC64;

Query Match	71.1%;	Score 4209.5;	DB 10;	Length 1088;
Best Local Similarity	70.2%;	Pred. No. 0;		
Matches 769;	Conservative 127;	Mismatches 179;	Indels 21;	Gaps 10;
QY	1	MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCOICGDDVGLAPGDPFVACNE	60	
Db	1	MNTGRLIAGSHNRNEFVLINADDTARIRSAEELSQTCKICRDEIELTDNGEPFIACNE	60	
QY	61	CAFPVCRDCYEREREGTQNCPOCKTRYKRLKGCORVTGDEEEDGVDDLDFNEFNWDGHS	120	
Db	61	CAFPTRCPCYEYERREGNQACPOCGTRYKRIKGSFVEGEDEEDDIDLEHEFY--GMDP	118	
QY	121	QSVAESML-YGHMSYGRGCBNGCAPQAFQNLNPNVPLLTNGQVDDIPPEQHALVPSFMG	179	
Db	119	EHVTEALYYMRLNTGRGTDEVSHLYSASPGSEVPLLTICDESDSMYSDRHALIVPSTG	178	
QY	180	GGRIRHPLPYADPSLPVQPSMDPSKDLAAYGYGSVAMKERENWKORQ-ERM---HOT	234	
Db	179	LGNRVHVPFTDSFASIHTRPMVPQKDLTVYGYGSVAMKDRMEVWKQOIIEKLQVKNER	238	
QY	235	GNDGG-----DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFPHY	289	
Db	239	VNDGGDGFIVDELDDPGLPMDEGRQPLSRKLPIRSSRINPYRMILFCRLAILGLFFHY	298	
QY	290	RVMHPVNDAFALWLISVCEIWFAMSWILDOFPKWFPIEREYLDRLSLRFDKEGQPSOL	349	
Db	299	RILHPVNDAFGLWLTSVCEIWFAMSWILDOFPKWPYIERETYLDRLSLRYPEKEGPSEL	358	
QY	350	APIDFVSTVDPLKEPPLVTNTVLSLSVDYPVDKVSQVYSDGGAAMLTFEALSETSEF	409	
Db	359	APVDVFSTVDPLKEPPLITANTVLSILAVDYPVEKAVACVYSDGGAAMLTFEALSYTAEF	418	
QY	410	AKKAVPECKRYNIEPRAPENYFOQKIDYLDKQVAANFVRERRAMKREYEEFKVRINALVA	469	
Db	419	ARKWYFCKKFSIEPRAPENYFSQKMDYLKHKVDPAFVMERRAMKRDYEEFKVINALVS	478	
QY	470	KAQKVPBEGWTMODGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRP	529	
Db	479	VSQKVPEDGWTMODGTPWPGNNVRDHPGMIQVFLGHSQVCDMDGNELPRLVYVSREKRP	538	
QY	530	YNHHKAGAMNALVRVSAVLTNAAYLLNLDCDHYINNSKAIKEACFMMDPLGKVCYV	589	
Db	539	FDHKKAGAMNSLIRVSAVLSNAPYLLNVDCHYINNSKAIREAMCFMMDPQSGKKICYV	598	
QY	590	QFPQRFDGIDRHDYANRNVVFFDINMKGLDGIIQPIYVGTGCVFRQALYGYDAPKTK	649	
Db	599	QFPQRFDGIDRHDYANRNVVFFDINMKGLDGIIQPIYVGTGCVFRQALYGFADAPKKQ	658	

QY	650	PPSR	T	C	N	C	M	P	K	M	C	F	C	C	C	C	F	G	N	R	K	O	K	T	T	P	K	T	E	K	K	L	L	F	K	E	N	O	S	P	A	L	G	I	D	E	A	P	709						
Db	659	PPGR	T	C	N	C	M	P	K	M	C	-	C	I	C	-	G	M	R	K	-	K	T	G	K	V	K	D	N	O	R	-	-	-	-	K	P	E	T	S	K	O	I	H	A	L	E	H	E	B	G	L	Q	711	
QY	710	GAEN	E	K	A	G	I	V	N	O	O	K	L	E	K	K	F	G	O	S	S	V	F	T	S	T	L	L	E	N	G	T	L	K	S	A	S	P	A	L	K	E	A	I	H	I	S	C	G	E	D	K	769		
Db	712	VTNA	E	N	N	S	E	T	A	Q	L	K	E	K	K	F	G	O	S	P	V	L	V	A	S	T	L	L	L	N	G	V	P	S	N	V	N	P	A	S	L	R	E	S	I	O	V	I	S	C	G	E	E	K	771
QY	770	TDMG	E	I	G	M	I	G	S	V	T	E	D	I	L	T	G	F	K	M	C	H	G	M	R	S	I	C	I	P	K	R	V	A	F	K	S	A	P	L	N	L	S	D	R	L	H	O	V	L	R	W	A	829	
Db	772	TEWG	K	E	I	G	M	I	G	S	V	T	E	D	I	L	T	G	F	K	M	C	H	G	M	R	S	V	C	M	P	K	R	A	F	K	S	A	P	I	N	L	S	D	R	L	H	O	V	L	R	W	A	831	
QY	830	IGSIE	I	F	F	S	N	H	C	P	L	M	Y	G	G	G	L	K	F	L	E	R	F	S	Y	I	N	S	I	V	P	W	T	S	I	P	L	A	Y	C	T	L	P	A	I	C	L	T	G	K	E	I	889		
Db	832	IGSVE	I	F	L	S	R	H	C	P	I	M	Y	G	G	G	L	K	M	L	E	R	F	S	Y	I	N	S	V	V	P	W	T	S	I	P	L	V	C	S	L	P	A	I	C	L	T	G	K	E	I	891			
QY	890	TPEL	N	N	V	A	S	L	M	F	M	S	L	F	I	C	I	F	A	T	S	I	L	E	M	R	M	S	G	V	I	D	M	W	R	N	E	Q	F	M	V	I	G	V	S	S	H	L	F	A	V	E	O	L	949
Db	892	VPEI	S	N	Y	A	G	I	L	F	L	M	F	M	S	I	A	V	T	G	I	L	E	M	O	M	G	K	I	G	I	D	M	W	R	N	E	Q	F	M	V	I	G	V	S	S	H	L	F	A	L	E	O	L	951
QY	950	LKVI	A	G	V	D	T	S	F	T	V	T	S	K	G	D	D	E	E	F	S	E	L	Y	T	F	K	W	T	L	L	P	T	L	L	L	N	F	I	G	V	A	G	S	N	A	I	N	G	1009					
Db	952	LKVL	A	G	V	S	T	N	F	T	V	T	S	K	A	D	D	E	F	S	E	L	Y	I	F	K	W	T	S	L	I	P	T	L	L	I	N	I	N	I	V	G	V	I	G	V	S	D	A	I	N	G	1011		
QY	1010	YESW	G	P	L	F	G	K	L	F	F	A	F	W	I	V	H	L	Y	P	L	K	G	L	V	G	R	O	N	R	T	P	T	I	V	I	W	S	I	L	L	A	S	I	F	S	L	M	W	R	I	D	F	1069	
Db	1012	YDSW	G	P	L	F	G	R	L	F	F	A	L	M	V	I	V	H	L	Y	P	L	K	G	L	L	G	O	D	R	V	P	T	I	I	L	W	S	I	L	L	A	S	I	L	T	L	M	W	R	V	N	P	1071	
QY	1070	LAKD	G	P	L																																																		

```

RESULT 8
O9FIB9
ID O9FIB9 PRELIMINARY; PRT; 1069 AA.
AC O9FIB9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Cellulose synthase catalytic subunit.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=99156233; PubMed=10048488;
RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
RT Sequence features of the regions of 1,081,958 bp covered by seventeen
RT physically assigned P1 and TAC clones.";
RT DNA Res. 5:379-391(1998).
DR EMBL: AB016893; BAB09408.1; -.
DR InterPro: IPR005150; Cellulose_synt.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synt; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PSS0089; ZF_RING_2; 1.
SQ SEQUENCE 1069 AA; 120861 MW; D03ED5C578DB3E7C CRC64;

```

Query Match	71.0%;	Score 4204.5;	DB 10;	Length 1069;
Best Local Similarity	70.6%;	Pred. No. 0;		
Matches 770;	Conservative 126;	Mismatches 166;	Indels 29;	Gaps 10;

[illegible]

Db	61	CAFPVCRPCYEYERREGNOSCQCKTRYKRIKGSPRVEGDEEDDGDIDLDPEFDY----	S	116
Qy	121	QSVAESMLYG--HMSYGRGDENGAPQAFQLNPNVPLLTNGQWDDIPPEQHALLVPSFMG		178
Db	117	RSGLSESEFSRRNSEFDLASAPGS-----QIPLLTYGEEDEVEISSDHALLIVSPSP		168
Qy	179	GGGKRIHPLPYADPSLPVQPRSMDP SKDLAAYGYGSVAWKERMENWKORERHQT--G		235
Db	169	GHIHRVQHPHPPDA--AHPRPVPOKDLAVGYGSVAWKDRMEEWKRQONEXYQVVKHD		226
Qy	236	NDGGDDGDADADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVYLGFPHYRVMHPV		295
Db	227	GDSLSGDGDADADIPMMEDEGRQPLSRKVPPIKSSKINBYRMLIVLRVLIGLFFHYRILHPV		286
Qy	296	NDAFALMLISVICEIWFAMSNLIDQFPKWPPIERETYLDRLSLRFQKEGQPSOLAPIDPF		355
Db	287	NDAYALMLISVICEIWFASVWLIDQFPKWPIERETYLDRLSLRYEKEGKPSLAGVDVF		346
Qy	356	VSTVDPLEKPPPLVTNTVLSLSVDYPVDKVSICYSDDGAAMLTFEALSETSEFAKKWP		415
Db	347	VSTVDPKKEPPLITANTVLSILAVDYPVDRVACIYSDDGAAMLTFEALSETAEFARKWP		406
Qy	416	FCRKRYNEPRAPEMYFOQKIDYLDKVAANFVRERAMKREYEEFKVRINALVAKAQVP		475
Db	407	FCCKYTIETPRAPEMYFCHKMDYLKNKVHAPAFVRERAMKRDYEEFKVINALVATAQKV		466
Qy	476	EEGWTMDGTWPMPGNNVDRHPMIQVFLGQSGGLDEGNEPLRLVYVSREKRPGYNHKK		535
Db	467	EEGWTMDGTWPMPGNNVDRHGMIOVFLGNNGVRDYENNELRLVYVSREKRPBGFDHKK		526
Qy	536	AGAMNALVRVSAVLTNAPYLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQPORF		595
Db	527	AGAMNSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPOSGKKICYVQPORF		586
Qy	596	DGIDRHRANRNVVFFDINMKGLDGIQGPITYGTGCVFRQALYGYDAPRTKKPPSRTC		655
Db	587	DGIDKSDRYSNRNVVFFDINMKGLDGLQGPITYGTGCVFRQALYGFDAPKKTKRMTC		646
Qy	656	NCWPKMCFCCCCFGNRKOKTKTPKTEKKLLFFKKEENQSPAYALGEIDEAPGAEN-E		714
Db	647	NCWPKMCLFCC--GLRKNRKS--KTDDKK---KNREASKQIHALENIEEGTKGTNDAA		698
Qy	715	KAGIVNQKLEKKFKGQSSVFVSTLLENGTLLKSASPASLKEAIIHVISCGYEDKTDMGK		774
Db	699	KSPAEOQLKEKKFKGQSPVFVASAGMENGLARNSPASLLREAIQVISCYEDKTEWGK		758
Qy	775	EIGWYGSVTEIDLTFQKHCHGWRSIYCI PKRVAFKGSAPLNLSDRLHOVLRWALGSIE		834
Db	759	EIGWYGSVTEIDLTFQKHSHGWRSVYCTPKIPAFKGSAPINLSDRLHOVLRWALGSVE		818
Qy	835	IFFSNHCPLMYGGGLKFLERFSYINSIVYPTWSLPLLAYCTLPALCILTGFKITPELN		894
Db	819	IFLSRHCPIWYGGGLKWLERLSYINSVVPWTSIPLLVYCSLPAICLTGKFIIVEIS		878
Qy	895	NVASLWENSLFICIFATSILENRMSGVGIIDWWRNEQFWVIGVSSHLPVFOGLLKVIA		954
Db	879	NYASILFMAIFGSIAVTGILEMQMGKVGIDDWWRNEQFWVIGVSAHLFALFOGLLKVIA		938
Qy	955	GVDTSFTVTSKGGDDEEFSSELTFFKTTLLIPPTLLLLNFIGVAVGSNAINNGYESWG		1014
Db	939	GVEITNFTVTSKAADDEGFSELYIFKMTSLLIPPTLLIINIVIGIVGISDAISNGYDSWG		998
Qy	1015	PLFGKLFFAFWVIVHLYPFLKGLVGRQNRTPPTIVIVMSILLASIFSLMWRIDPFLAKD		1074
Db	999	PLFGRLFFFAFWVILHLYPFLKGLLGQDRMPTIILVMSILLASITLLMVRVNPVFAK-G		1057
Qy	1075	GPLLECGCLDC	1085	
Db	1058	GPILBICGLDC	1068	

## RESULT 9

065338  
ID 065338 PRELIMINARY; PRT; 1081 AA.  
AC 065338;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE Cellulose synthase (Fragment).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Wu L., Joshl C.P., Chiang V.L.;  
RT "AraxcelA, a new member of the cellulose synthase gene family from  
RT Arabidopsis (Accession No. AF062485) (PGR98-114).";  
RL Plant Physiol. 117:1125-1125(1998).  
DR EMBL; AF062485; AAC29067.1; -.  
DR InterPro; IPR005150; Cellulose\_synth.  
DR InterPro; IPR001841; Znf ring.  
DR Pfam; PF03552; Cellulose\_synth; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PSS0089; ZF\_RING\_2; 1.  
FT NON TER 1  
SQ SEQUENCE 1081 AA; 122446 MW; DC59A35A1713FD9F CRC64;

Query Match 71.0%; Score 4202; DB 10; Length 1081;  
Best Local Similarity 70.8%; Pred. No. 0;  
Matches 770; Conservative 126; Mismatches 172; Indels 20; Gaps 10;

QY 7 LVAGSHNRNELVIRRDGDPKPREQNGQVQICGDDVGLAPGDPFVACNECAFVVC 66  
DB 4 LTAGSHNRNEFLVINADENARIRSVQELSGQTCQICRDEIELTVDEPFVACNECAFVVC 63  
QY 67 RDCYERRECTONCPQCKTRKYKRLGQCVTGDDEEDGVDDLNEFNW--DGHDSQSV 124  
DB 64 RPYEYERREGNQACPQCKTRFKRLKSPVEGDEEDDIDLDNEFEYGNNGIGFDYVS 123  
QY 125 ESMLYGHMSYG-RGGDPNGAPQAFQNLNPNVPLLTNGQWVDIPEQHAL-VPSFMGGGK 182  
DB 124 EGMGISRRNSGFPQSDLDSPGSGQ---IPLLTYGDEDEVEISSDRHALIVPPLSGHGN 179  
QY 183 RIHPLPYADPSLPVQPRSMPSKDLAAYGYSVAKERMENWKOR-ERMHOTGNDGGD 241  
DB 180 RVHPVSLSDPTVAHRRLMVPQKDLAVYGYGSVAMKDRMEEMKQNEKIQVVRHEGDPD 239  
QY 242 --DGDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVGFFFHYRVMPVNDAP 299  
DB 240 FEDGDADFPMDDEGRQPLSMKIPKSSKINPYRMLIVRLVILGLFHYRILHPVKDAY 299  
QY 300 ALWLISVICEIWFAMSWILDQFPKMFPIERTETYLDRLSLRFCKEQPSOLAPIDFVSTV 359  
DB 300 ALWLISVICEIWFVAVSWLDQFPKMFPIERTETYLDRLSLRFCKEQPSGLSPDVFTV 359  
QY 360 DPLKEPPLTNTVLSILSVDPVVKVSCYVSDGAAMLTFEALSETSEFAKWPFPCKR 419  
DB 360 DPLKEPPLTANTVLSILAVDPVDKACVYSDGAAMLTFEALSETAEFARKWVPFCKK 419  
QY 420 YNIEPRAPBWYFOQKIDYLDKQVAANFVRRERAMKREYEEFKVIRINALVAKAQVPEEGW 479  
DB 420 YCIEPRAPBWYFCHKMDYLDKQVAPFRRERAMKRDYEEFKVIRINALVATAQKVPEDGW 479  
QY 480 TMQDGTWPGNNVRDHPGMIQVFLGSGGLDCEGNELPRLVYVSREKRPGYNHKKAGAM 539  
DB 480 TMQDGTWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPFGDHHKKAGAM 539  
QY 540 NATVRVSAVLTNAPYLLNDCDHYIINNSKAIKEAMCFMMDPLGKKYCYVQFPQRFDGID 599  
DB 540 NSLIRVSGVLSNAPYLLNVDCHYIINNSKALREAMCFMMDPQSGKKICYVQFPQRFDGID 599  
QY 600 RHDRYANRVVFFDINMKGLDIGPIYVGTGCVFRQALYGYDAPKTKKPSRTCNCPW 659

DB 600 RHDRYSNRNVFFDINMKGLDLQGPITYGTGCVFRQALYGFADAPKTKKGPRTKCNCPW 659  
QY 660 KWCFCCCCFCGNRKQKTTKPKTEKKLLFPKKEENQSPAYALGEIDEA--PGAENEKAG 717  
DB 660 KWCCL--LCFGSRKRRKAKTVADKK---KNREASKQIHALENIEEGRGHKVLNVEQST 713  
QY 718 IVNQKLEKFKGSSVFVTSTLLENGTLLKSASPASLLKEAIVHISGVEDKTDWKEIG 777  
DB 714 EAMQMKLQKKYGGSSVFVASARLENGMARNASPACLLKEAIVISRGVEDKTEWKEIG 773  
QY 778 WIYGSVTEDILTGFKHCHGMRISYICIPKRVAFKGSAPLNLSDRLHQVLRWALGSIIEIF 837  
DB 774 WIYGSVTEDILTGSKMHSHGMRHVCTPKLAAFKGSAPINLSDRLHQVLRWALGSEIIFL 833  
QY 838 SNHCPPLWYGGGLKFLERFSYINSIVYPTWISIPPLAYCTLPALICLTGKFTPELNVA 897  
DB 834 SRHCPWYGGGGLKFLERLSYINSVYPTWISIPPLIVYCSLPALICLTGKFTVPEISNYA 893  
QY 898 SLWFMSLFICIFATSILEMWRSVGIDDMWRNEQFWVIGVSSHLPFAVQGLLKVIAGVD 957  
DB 894 SILFMALFSSIAITGILEMQGWKVGIDDMWRNEQFWVIGVSAHLFALQGLLKVLAVGD 953  
QY 958 TSFTVTSKGGDDEERSSELYTEKWTLLIPPTLLILNFIGVAVGSNAINNGYESWGPIF 1017  
DB 954 TNFTVTSKAADDEERSDLYLFKWTSLIPPTLLIINVIGVAVGSDAISNGYDSWGPIF 1013  
QY 1018 GKLFFAFVIVHLYPELKLGVGRQNRTPTIIVMSILLASIFSLWVRIDPELAKDDGPL 1077  
DB 1014 GRLFFALWVITHLYPELKLGLLGKQDRMPTIIVWSILLASITLLWVRVNFVAK-GGPI 1072  
QY 1078 LEECGLDC 1085  
DB 1073 LEICGLDC 1080

RESULT 10

Q9LLI9  
ID Q9LLI9 PRELIMINARY; PRT; 1075 AA.  
AC Q9LLI9;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Cellulose synthase-1.  
GN CESA-1.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoidae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20398328; PubMed=10938350;  
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,  
RA Xocoostle-Cazares B., Delmer D.P.;  
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene  
RT family";  
RL Plant Physiol. 123:1313-1324(2000).  
DR EMBL; AF200525; AAF89961.1; -.  
DR InterPro; IPR005150; Cellulose\_synth.  
DR InterPro; IPR001841; Znf ring.  
DR Pfam; PF03552; Cellulose\_synth; 1.  
DR SMART; SM00184; RING; 1.  
SQ SEQUENCE 1075 AA; 121181 MW; 67A4FBD97A811F33 CRC64;

Query Match 67.3%; Score 3984; DB 10; Length 1075;  
Best Local Similarity 66.4%; Pred. No. 0;  
Matches 738; Conservative 143; Mismatches 167; Indels 64; Gaps 15;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGQVQICGDDVGLAPGDPFVAC 58  
DB 1 MAANKGMVAGSHNRNEFVMIRHDGDPGSAKPTYSANGQVQICGDSVGSATGDFVAC 60





```
QY 577 MMDPLGKVCYVQFPQRPDGDIDRHRANRVNVPDINMGLDGIQGPITYGTGCVFRR 636
Db 574 MMDPALGRKTCYVQFPQRPDGDIDLHRYANRNIVFPDINMGLDGIQGPITYGTGCCFNR 633
QY 637 QALYGYDAPKTKPPSRCTNCWPKWCFCCECFGNRKOKTTKPKTEKKLLFFKKEENOS 696
Db 634 QALGYD-----FVLTEADLEPNIVKSCC-GRKRKNKSYMSQSR--IMKRTESSA 683
QY 697 PAYALGEIDEAAGAENEKAGIVNOQKLEKFGQSSVFVSTLLENGTLKSASPASLTK 756
Db 684 PIFNMEDIEEGIEGVEDERSVLSQRKLEKRGOSPIFIASFTMTGGIPSTNPASLTK 743
QY 757 EAIHVISCYEDKTDWCKEIGWYGSVTEIDLTFKMHCHGWRISIYCIIPKRAFKGSAPL 816
Db 744 EAIHVISCYEDKTEWCKEIGWYGSVTEIDLTFKMHARGWOSIYCMPPRPFCKGSAPI 803
QY 817 NLSDRHLQVLRWALGSIIEIFFSNHCPLWYGGGLKFLERFSYINSIYVWTSIPLAYC 876
Db 804 NLSDRNLQVLRWALGSIIEILSRHCPIWYGNRLKLERLAYINTIYVITSVPLIAYC 863
QY 877 TLPAICLLTGKFTPELNNVASLWMSLFCIFATSILEMRWSGVGIDWWRNEQFWVIG 936
Db 864 VLPAILCLTNKFIPEISNYAGMFFILLFASIFATGILELRWSGVGIEDWWRNEQFWVIG 923
QY 937 GVSSHLFAVFOGLKVIAGVDTSFTVTSKGD-DEEFSELYTFKWTLLIPPTLLILNF 995
Db 924 GTSALHFAVFOGLKVLGIDTNTFTVTSKASDEGDFAELVFKWTSLLIPPTVLVINL 983
QY 996 IGVNAGVSNAINNGYESWGPFLFGKLFPAFWIVHLYPFLKGLVGRQNRPTIIVWSILL 1055
Db 984 VGMVAGISYAINSGYQSWGFLFGKLFPSIWLHLYPFLKGLMGRQNRPTIIVWSILL 1043
QY 1056 ASIFSLWVRIDPFLA-KDDGPLEECGLDC 1085
Db 1044 ASIFSLWVKIDPFIPTQKAAALGQGVNC 1074
RESULT 12
Q9SWW6 PRELIMINARY: PRT; 1026 AA.
ID Q9SWW6;
AC Q9SWW6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Cellulose synthase catalytic subunit (AT5g17420/T10B6_80).
GN IRX3 OR T10B6_80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. LANDSBERG RECTA;
RX MEDLINE=99264300; PubMed=10330464;
RA Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
synthase required for secondary cell wall synthesis.";
RL Plant Cell 11:769-780(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
```

```
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Saito M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender B.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinzaki K., Davis R.W., Theologis A.,
RA Becker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF091713; AAD4085.1; -
DR EMBL; AL391142; CAC01737.1; -
DR EMBL; AY139754; AAM98075.1; -
DR InterPro; IPR005150; Cellulose_synt.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1026 AA; 115797 MW; 503BFC78BE6E511 CRC64;

Query Match 67.0%; Score 3967.5; DB 10; Length 1026;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 736; Conservative 129; Mismatches 145; Indels 91; Gaps 16;

QY 1 MEASAGLVAGSHNRNELVIRDDDPGPKPREQNGVQICGDDVGLAPGDPFVACNE 60
Db 1 MEASAGLVAGSHNRNELVIRHNEE--PKPLKMLDQGFCEICGDIQGLTVEGDLFVACNE 58
QY 61 CAFPVCRDCYEYERREGTONCPCKTRYKRLKGCORVTGDEEEDGVDDLNEFNWDGHS 120
Db 59 CGFPACRDCYEYERREGTONCPCKTRYKRLRGSPRVEGDEDEDIDIEYEFNIE-HEQ 117
QY 121 ---QSVAESMLYGHMSYGRG--GDPNGA-PAFQLNPNVPLLTNGQVDDIPPEQHAIYP 174
Db 118 DKHKHSAAEAMLYGKMSYGRGPEDDENGRFP-----PVLAGHSGE----- 157
QY 175 SFGGGG-----KRIHPLPYADPSLPVQPRSMPSKDLAAYGYSVAMKERMENWK 225
Db 158 -FPVGGGYGNGEHLKRVHPY-----PSSKAGSEG---GWRERMDWK 197
QY 226 QRQERMGQTNDGGCGDDGDDADLPMDERQQLSRKIPLPSSQINPYRMIIIRLVVLGF 285
Db 198 L-----QGNLGPEDD-DDPEMGLDEARQPLSRKVPFIASSKINPYRMIVARLVILAV 250
QY 286 FPHYRVMPVNDAPALMLISVICETWPFAMSWILDQFPKWPPIERETYLDRLSRFDEKQ 345
Db 251 FLRYRLNPVHDALGLWLTSVICETWFAVSWILDQFPKWPPIERETYLDRLSRYEREGE 310
QY 346 PSQ LAPIDFFVSTVDPLKEPPLVTNTVLSISVDYFVDKVCYVSDGAAMLTFEALSE 405
Db 311 PNM LAPVDFVSTVDPLKEPPLVTSNTVLSILAMDYFVEKISCVSDGASMLTFEESLSE 370
QY 406 TSEFAKKWVPECKRYNIEPRAPENYFQOKIDYLDKVAANFVRBRAMKREYEEFKVRIN 465
Db 371 TAEFARKWVPECKKFSIEPRAPENYFTLKVDYLQDKVHPTFYERRRAMKREYEEFKVRIN 430
QY 466 ALVAKAQVVEEGWTMODGTPWPGNNVRDHPGMIQVFLGSGGLDCEGNELPRLVYSRE 525
Db 431 AQVAKASKVPLEGIMODGTPWPGNNTKDPGMIQVFLGHSGGFDVEGHELPRLVYSRE 490
QY 526 KRPGYNHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKK 585
Db 491 KRPGFQHHKAGAMNALVRVAGVLTNAPFMLNLDCHYVNNSKAVREAMCFLMDPQIGKK 550
QY 586 VCYVQFPQRPDGDIDRHRANRVNVPDINMGLDGIQGPITYGTGCVFRRQALYGYDAP 645
Db 551 VCYVQFPQRPDGDIDTNDRYANRVNVPDINMGLDGIQGPITYGTGCVFRRQALYGYEP 610
QY 646 KTKKPPSRCTNCWPKWCFCCECFGNRKOKTTKPKTEKKLLFFKKEENOSPAYALGEID 705
Db 611 KGPKRPKMIS-----CGCCPCFGRRRNKK-----FSKNDMNGDVAALG--- 649
QY 706 EAAPGAENEKAGIVNOQKLEKFGQSSVFVSTLLENGTLKSASPASLTKKAIHVISCG 765
Db 650 ----GAEQDEHLMSEMFECTFGQSSIFVTSTLMEGGVPPSSSPAVLLKAIHVISCG 705
QY 766 YEDKTDWCKEIGWYGSVTEIDLTFKMHCHGWRISIYCIIPKRAVFKGSAPLNLSDRLHQV 825
```

Db 706 YEDKTEWGTGLWIMYGSITEDILTGFKMHCGRWSIYCMKRPAPFGSAPINLSDRLNQV 765  
Qy 826 LRMALGSIEIFPSNHCPLWYGY-GGGLKFLERFSYINSIVYPTSIPLAYCTLPALCL 884  
Db 766 LRMALGSVEIFFSRHSPLWYGKGLKMLERFAYANTTIYPTISIPLAYCILPAICLL 825  
Qy 885 TGKFTPELNNVASLWMSLFICIFATSILEMWMSGVGIDDMWRNEQFWVIGVSSHLEFA 944  
Db 826 TDKFIMPPISTFASLFFISLFMSIIVTGILELRWSGVSIEMWRNEQFWVIGISAHLEFA 885  
Qy 945 VFQGLLKVIAGVDTSFVTSSKGGDEEFSSELYTFKMTLLIPPTLLLLNFIGVAVGSN 1004  
Db 886 VVQGLKLKILAGIDTNFTVTSKATDDDDFGEIYAFKWTLLIPPTVLIINIVGVAGISD 945  
Qy 1005 AINNGYESWGPLFGKLFPAFWIVHLYPFLKGLVGRQNRPTTIIVWSILLASISFLMW 1064  
Db 946 AINNGYQSWGPLFGKLFSSFWIVHLYPFLKGLMGRQNRPTTIIVWSVLLASISFLMW 1005  
Qy 1065 RIDPFLAKDDGPLLEECGLDC 1085  
Db 1006 RIDPFVLKTKGPDTSKCGINC 1026

RESULT 13

Q9XHP6 PRELIMINARY; PRT; 1026 AA.  
ID Q9XHP6  
AC Q9XHP6;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Cellulose synthase catalytic subunit.  
GN IRX3.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Landsberg erecta;  
RX MEDLINE=99264300; PubMed=10330464;  
RT Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;  
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose  
RT synthase required for secondary cell wall synthesis.";  
RL Plant Cell 11:769-780(1999).  
DR EMBL; AF088917; AAD32031.1; -.  
DR InterPro; IPR005150; Cellulose\_synth.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF03552; Cellulose\_synth; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
SQ SEQUENCE 1026 AA; 115858 MW; 453BFD1D283C4D70 CRC64;

Query Match 67.0%; Score 3963.5; DB 10; Length 1026;  
Best Local Similarity 66.8%; Pred. No. 0;  
Matches 736; Conservative 128; Mismatches 146; Indels 91; Gaps 16;

Qy 1 MEASAGLVAGSHNRNELVIRRDGDPKPRREQNGQVCQICGDDVGLAPGDDPFVACNE 60  
Db 1 MEASAGLVAGSHNRNELVIRHNEE--PKLKNLDGQFCEICGDQIGLTVEGDLFVACNE 58  
Qy 61 CAFPVCRDCEYERREGTQNCPOCKTRYKLGKCQRTVTGDEEEDGVDDLNEFMWDGHS 120  
Db 59 CGFPACRPCYERERREGTQNCPOCKTRYKLGSPRVEGEDEDEEDIDIEYEFNIE-HEQ 117  
Qy 121 ---QSVAESMLYGHMSYGRG-GDPNGA-PQAFQNLNPNVPLLTNGQVDDIPPEQHALVP 174  
Db 118 DKHKHSAEAMLYGKMSYGRGPEDDENGFRP-----PVIAGHSGE----- 157  
Qy 175 SFMGGGG-----KRHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAWKERMENWK 225  
Db 158 -FPVGGGYGNGEHGLHKRVHPY-----PSSEAGSGE-----GWRERMDDWK 197

Qy 226 QROERHQTGNDGGDDGDADLPLMDEARQOLSRKIPSSQINPYRMIIIRLVYLG 285  
Db 198 L-----QHGNLGPEDD-DBDEMGLIDEARQPLSRKVP IASSKINPYRMIVARVLAV 250  
Qy 286 FHHRYVMHPVNDALFALWLTISVCEIWFAMSWITLDQFPKMPPIERETYLDRLSRFDEGQ 345  
Db 251 FLRYRLNLPVHDALGLWLTISVCEIWFAMSWITLDQFPKMPPIERETYLDRLSRYEREGE 310  
Qy 346 PSQIAPIDFVSTVDPLKEPPLVTNTVLSLSVDYVDKVCYVSDGAAMTFPAISE 405  
Db 311 PNMILAPVDVFSVTDPLKEPPLVTSNTVLSILAMDYPERKISCVSDDGASMLTFESISE 370  
Qy 406 TSEFAKMWPFCKRYNIEPRABEWYFQOKIDYLLKDKVAANFVRERRAMKREYEEFKYRN 465  
Db 371 TAEFAKMWPFCKKFSIEPRABEMYFTLKVDYLDQKHPTFYKERAMKREYEEFKYRN 430  
Qy 466 ALVAKAQKVPBEGWTMDGTPWPGNNVRDHPGMIQVFLQSGGLDCEGNEPLRLVYVRE 525  
Db 431 AQVAKASKVPLEGIMODGTPWPGNNTKDHPGMIQVFLHSGGFDVEGHELPRLVYVRE 490  
Qy 526 KRPGYNHHKAGAMNALVRVSAVLTNAPYLLNLDCHYTNNSKAIKBAKMPMDPLGKK 585  
Db 491 KRPGFOHHKAGAMNALVRVAGVLTNAPFMLNLDCDHYNNNSKAVREAMCFMLDPQIGKK 550  
Qy 586 VCYVQFPQRFQDIDRHRYANRNVFPDINMKGLDGIGPIYVGTGCVFRQALYGDAP 645  
Db 551 VCYVQFPQRFQDIDTNDRYANRNTVFPDINMKGLDGIGPYVGTGCVFRQALYGYEPP 610  
Qy 646 KTKKPPSRTCNCPKWCFCGCCGFGNRKQKTTKPKTEKKLLFFKKEENOSPAYALGETD 705  
Db 611 KGPKRPKWIS-----CGCCPCFGRRKMKK-----FSKNDMNGDVAALG--- 649  
Qy 706 BAAPGAENEKAGIVNQKLEKKGSSVFVTSTLLENGTLKASAPASLLKEAIVHISCG 765  
Db 650 ---GAEKDKEHLMFEMNFEKTFGQSSIFVTSTLMEBGGPSPSSPAVLLKEAIVHISCG 705  
Qy 766 YEDKTDWKEIGWYIGSVTEDILTGFKMHCGRWSIYCTPKRVAFKGSAPINLSDRLNQV 825  
Db 706 YEDKTEWGTGLWIMYGSITEDILTGFKMHCGRWSIYCMKRPAPFGSAPINLSDRLNQV 765  
Qy 826 LRMALGSIEIFPSNHCPLWYGY-GGGLKFLERFSYINSIVYPTSIPLAYCTLPALCL 884  
Db 766 LRMALGSVEIFFSRHSPLWYGKGLKMLERFAYANTTIYPTISIPLAYCILPAICLL 825  
Qy 885 TGKFTPELNNVASLWMSLFICIFATSILEMWMSGVGIDDMWRNEQFWVIGVSSHLEFA 944  
Db 826 TDKFIMPPISTFASLFFISLFMSIIVTGILELRWSGVSIEMWRNEQFWVIGISAHLEFA 885  
Qy 945 VFQGLLKVIAGVDTSFVTSSKGGDEEFSSELYTFKMTLLIPPTLLLLNFIGVAVGSN 1004  
Db 886 VVQGLKLKILAGIDTNFTVTSKATDDDDFGEIYAFKWTLLIPPTVLIINIVGVAGISD 945  
Qy 1005 AINNGYESWGPLFGKLFPAFWIVHLYPFLKGLVGRQNRPTTIIVWSILLASISFLMW 1064  
Db 946 AINNGYQSWGPLFGKLFSSFWIVHLYPFLKGLMGRQNRPTTIIVWSVLLASISFLMW 1005  
Qy 1065 RIDPFLAKDDGPLLEECGLDC 1085  
Db 1006 RIDPFVLKTKGPDTSKCGINC 1026

RESULT 14

Q8GSW2 PRELIMINARY; PRT; 1032 AA.  
ID Q8GSW2  
AC Q8GSW2;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Cellulose synthase.  
GN CESA2.  
OS Populus tremuloides (Quaking aspen).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Salicaceae; Populus.  
OX NCBI\_TaxID=3693;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Xylem;  
RX MEDLINE=22271539; PubMed=12383501;  
RA Samuga A., Joshi C.P.;  
RT "A new cellulose synthase gene (PtcCesA2) from aspen xylem is  
RT orthologous to Arabidopsis AtCesA7 (irx3) gene associated with  
RT secondary cell wall synthesis.";  
RL Gene 296:37-44(2002).  
DR EMBL; AY095297; AAM2629.1;  
SQ SEQUENCE 1032 AA; 116958 MW; 15ABB28FDC60B96D CRC64;  
  
Query Match 66.8%; Score 3955.5; DB 10; Length 1032;  
Best Local Similarity 66.4%; Pred. No. 0;  
Matches 729; Conservative 138; Mismatches 152; Indels 79; Gaps 15;  
  
QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPPREONGVQICGDDVGLADGDPFVACNE 60  
DB 1 MEASAGLVAGSHNRNELVIRH--GHEEHKPLKMLDQVCEICGDEIGLTVDDGLFVACNE 58  
  
QY 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGTDEEEDGVDDLDNEFNWDGHS 120  
DB 59 CGFPVCRPCYEYERREGTQNCPOCKTRYKRLKGSFVEGDDEDDVDDEHEFIIEQD 118  
  
QY 121 QS--VAESMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQVDDIPPEQHALVPSFMG 178  
DB 119 KNKHLTEAMLHGKMTYGRGHD-----EENSQFPFVITGIRSRPVSGE-----FSIG 165  
  
QY 179 GGG-----KRTHPLPYADPSLPVQPRSMDSKDLAAYGYGSVANKERMENMKQOE 229  
DB 166 SHGEQMLSSSLHKRVHPYVSEPG---SARWDEKE-----GSKERMDEWKM--- 210  
  
QY 230 RMHQTGNDGGDGDADPLMDARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFPHY 289  
DB 211 ---QHGNLGPED-DAAEAMLEDARQPLSRKVPASSKINPYRMVIVARLIILAVFLRY 266  
  
QY 290 RVMPVNDAAFALWLSVCEIWFAMSWILDQPKWPIERETYLDRLSLRDEKQGPSOL 349  
DB 267 RILHPVHDALGLWLTISVCEIWFALISWILDQPKWLPIDRETYLDRLSLRDEKQGPSNML 326  
  
QY 350 APIDFVSTVDPLEKPPVLTNTVLSTLSVDYPVDKVCYVSDGAAMLFREALSETSEF 409  
DB 327 APVDVFSVSTVDPKPEPLVTGNTLLSLAMDYPVEKISCYLSDGASMCFTREAMSETAEF 386  
  
QY 410 AKKWPFCRKNIEPRAPRWYFOQKIDYLDKQVAANFVRERAMKREYEERFVRINALVA 469  
DB 387 ARKWPFCRKNIEPRAPRWYFTLVKVDYLDKQVPTFKERRAMKREYEERFVRINALVA 446  
  
QY 470 KAQKVPEEGWTMDGTFWPGNNVRDHPGMIQVFLGSGGLDCEGNEPLRLVVSREKRP 529  
DB 447 KAQKVPEEGWTMDGTFWPGNNVRDHPGMIQVFLGSGGHDEGNEPLRLVVSREKRP 506  
  
QY 530 YNHKKAGAMNALVRSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLLKKVCYV 589  
DB 507 FSHKKAGAMNALIRVLAILTNAFMLNLDCDHYVNSKAVREAMCFMMDPLLKKVCYV 566  
  
QY 590 QFPQRFIDGIDRHRVYANRVVFFDINMKGLDGIGPIYVGTGCVFRRQALGYDAPK-TK 648  
DB 567 QFPQRFIDGIDRHRVYANRVVFFDINMKGLDGIGPIYVGTGCVFRRQALGYDAPKDPK 626  
  
QY 649 KPPSRTCNCPKWCFCGCCFGRNKKQKTKTKTEKKLLFFKKEENQSPAYALGEIDEAA 708  
DB 627 RPKMETCD-----CCPCFGRKKKNA---KT-----GAVVEGM 656  
  
QY 709 PGAENEKAGIVNQKLEKFGQSSVFVTSTLLENGTLLKSASPASLKEATHVISCYED 768  
DB 657 D--NNDKELLMSHNFEEKFGOSAI FVTSTLMEEGVPPSSSPALLKEATHVISCYED 714  
  
QY 769 KTDWGKEIGWIYGSVTEDILLTGFKHCHGWRSTYCIPIKRVAFKGSAPLNLSDRLHQVLRW 828  
DB 769 KTDWGKEIGWIYGSVTEDILLTGFKHCHGWRSTYCIPIKRVAFKGSAPLNLSDRLHQVLRW 828

DB 715 KTEWGLGWIYGSITEDILTGFKHCHGWRSTYCMKRAAFKGSAPINLSDRLNQVLRW 774  
QY 829 ALGSIETFFSNHCPWYGY-CGGLKFLERFSYINSIVPWTSLPLAYCTLPACLLTGK 887  
DB 775 ALGSVEIFFSGHSPNMGYKKGKLMLEFAYVNTTTPYPSLALVAYCCLPALCLLTDK 834  
QY 888 FITPELNNVASLWMSLFCIFATSILEMWMSGVGIDMWRNEQFWVIGVSSHLFAVQ 947  
DB 835 FIMPEISTFASLFFIALFLSIFSTIGLELRMSGVSIEMWRNEQFWVIGVSAHLFAVQ 894  
QY 948 GLKVIAGVDTSTFTVTSKGGDEEFSELYTFKWTLLIPPTLLLNFIGVAGVSNAIN 1007  
DB 895 GLKVIAGIDLNFTVTSKATDDDFGELYAFKWTLLIPPTLLINLVGAVAGSDAIN 954  
QY 1008 NGYESWGPLFGKLFPAFWIVHLYPFLKGLVGRQNRTPTIIVMSILASIFSLWVRID 1067  
DB 955 NGYQSWGPLFGKLFPAFWIVHLYPFLKGLMGRQNRTPTIIVMSVLLASIFSLWVRID 1014  
QY 1068 PFLAKDDGPLLEECGLDC 1085  
DB 1015 PFVMTKTRGPDTKQCGLNC 1032  
  
RESULT 15  
O48946 PRELIMINARY; PRT; 1081 AA.  
ID O48946  
AC O48946;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Cellulose synthase catalytic subunit.  
GN RSM1 OR F8B4.110 OR AT4G32410.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=98111412; PubMed=9445479;  
RA Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Herth W.,  
RA Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,  
RA Redmond J., Williamson R.E.;  
RT "Molecular analysis of cellulose biosynthesis in Arabidopsis.";  
RL Science 279:717-720(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Terry N., Ardiles W., Buysshaert C., Dasseville R.,  
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
RA Villaroel R., Gelen J., Van Montagu M., Hohenseel J., Mewes H.W.,  
RA Mayer K.F.X., Lemcke K., Schueller C.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Submitted sequencing project;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Terry N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,  
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,  
RA Gelen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Submitted sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF027172; AAC39334.1; -  
DR EMBL; AL034567; CAA22568.1; -  
DR EMBL; AL161581; CAB79958.1; -  
DR InterPro; IPR005150; Cellulose\_syn. InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF03552; Cellulose\_syn; 1.

DR SMART; SM00184; RING; 1;  
SQ SEQUENCE 1081 AA; 122236 MM; BDEB5D9DEE334D59 CRC64;

Query Match 66.7%; Score 3945.5; DB 10; Length 1081;  
Best Local Similarity 66.9%; Pred. No. 0;  
Matches 727; Conservative 140; Mismatches 182; Indels 37; Gaps 11;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPPRONGQVCQICGDDVGLAPGDPFVACNE 60  
DB 1 MEASAGLVAGSYRNLVIRRHESDGTGKPLKMNNGQICQICGDDVGLAETGDFVACNE 60  
QY 61 CAFVPCRCDCYEREREGTQNCPOCKTRKYKLGCCQRTVGDEBEDGVDDLNEFNMGDHDS 120  
DB 61 CAFVPCRCDCYEREREGTQNCPOCKTRKYKLGCCQRTVGDEBEDGVDDLNEFNMGDHDS 120  
QY 121 QSVAESMLYGHMSYGRGDDPENGAPQAFQNLNPNVPLLTNGQMVDD--TPPEQHALVPSF- 176  
DB 117 QGANKA-----RHQRHGEFEFSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTTSGP 169  
QY 177 MGGGGRIRHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAKEMENKQOER--MHQT 234  
DB 170 LGPSDRNAISSPYIDPRQVPVRIVDPSPKDLNSYGLGNVDWKEKRVGKWKQEKMLQMT 229  
QY 235 GN-----DGGDDGDDADLPLMDEARQOISRKIPLSSQINPYRMIIIRLVILGF 285  
DB 230 GKYHEGKGGEIEGTGSGNGE--ELQMAADTRLPMRSRVPIPPSSRLTPYRVIIIRLIIICF 287  
QY 286 FEHYRVMPVNDALFALMLISVCEIWFAMSNILDOFPKMFIERETYLDRLSLRFDEKQ 345  
DB 288 FLQYRTHPVGNAYPLMLTSVCEIWFAMSNILDOFPKMFIERETYLDRLSLRFDEKQ 347  
QY 346 PSQAPIDFVSTVDPLKEPPLVTNTVLSISVDYVPDKVSCYVSDGAMLTPEALSE 405  
DB 348 PSQLPVVDVFSVTDPLKEPPLVTANTVLSISVDYVPDKVACVSDGAMLTPEALSE 407  
QY 406 TSEFAKMWPFCKRYNIEBPAPWEYFOQKIDYLDKVAANFRERERAMKREYEEFKVRIN 465  
DB 408 TAEFAKMWPFCKRYNIEBPAPWEYFOQKIDYLDKVAANFRERERAMKREYEEFKVRIN 467  
QY 466 ALVAKAQKVPBEGWMTQDGTMPGNNVRDHPGMIQVFLGSGGLDCBGNELPRLVYVSRE 525  
DB 468 ALVAKAQKVPBEGWMTQDGTMPGNNVRDHPGMIQVFLGSGGLDCBGNELPRLVYVSRE 527  
QY 526 KRPYNNHHKAGANALVRSVAVLTNAPYLLNDCDHYINNSKAIKEAMCEMMDPLLKX 585  
DB 528 KRPYNNHHKAGANALVRSVAVLTNAPYLLNDCDHYINNSKAIKEAMCEMMDPLLKX 587  
QY 586 VCYVQFPQRFDDGIDRHDRYANRVVFPDINMKGLDGIQGPITYGTGCVFRQALYGYDAP 645  
DB 588 CCYVQFPQRFDDGIDRHDRYANRVVFPDINMKGLDGIQGPITYGTGCVFRQALYGYDAP 647  
QY 646 KTKPPSRRTCNCPKWCFCFCCCFGNRKQKTKTKTEKKLLFFKKEENQSPAYALGEID 705  
DB 648 LTEE-----DLEPNIIVKSCCGSRKKKSKSKYNYEKRRGI--NRSDSNAPLFNMEDID 699  
QY 706 EAPGAENKAGIVNOQKLEKFGQSSVFVTSTLLENGTLKASAPASLKEAIIHVISC 765  
DB 700 EGFEYDDERSILMSQSVKRFQSPVFIATFMEQGGIPTTNPATLKEAIIHVISC 759  
QY 766 YEDKTWGEIGWYGSVTEDILTFKMHCHGWSIYCIKRVAFKGSAPLNLSDRLHQV 825  
DB 760 YEDKTWGEIGWYGSVTEDILTFKMHCHGWSIYCIKRVAFKGSAPLNLSDRLHQV 819  
QY 826 LKVALGSEIIFPSNCPMLWYGGGLKFLERFSYINSIVYPTWSIPLAYCTLPALICLLT 885  
DB 820 LKVALGSEIIFPSNCPMLWYGGGLKFLERFSYINSIVYPTWSIPLAYCTLPALICLLT 879  
QY 886 GKFTPELNNVASLWMSLFCIPATSILEMWGSVIGIDDWNRNEQFWVIGVSSHLPFV 945  
DB 880 DRFTIPEISNYASIWFLFISIAVTGILBLRWGSVIEDWNRNEQFWVIGVSSHLPFV 939  
QY 946 FQGLLKVLAGVDTSTFTVTSKGD-DEEFSSELYTFKWTLLIPTTLLLNFIQVAVGSN 1004  
DB 946 FQGLLKVLAGVDTSTFTVTSKGD-DEEFSSELYTFKWTLLIPTTLLLNFIQVAVGSN 1004

DB 940 FQGLLKVLAGIDTNTFTVTSKATDEDDGPAELYIFKWTALLIPTTLLNLGIVAGVSY 999  
QY 1005 AINNGYESWGPLFGKLFFAFWIVHLYPFLKGLVGRQNRPTIVIVMSILASISFLMW 1064  
DB 1000 AVNSGYQSWGPLFGKLFALWVIAHLYPFLKGLGRQNRPTIVIVMSVLLASISFLMW 1059  
QY 1065 RIDPFL 1070  
DB 1060 RINPFV 1065

Search completed: December 15, 2003, 13:04:19  
Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:56:44 ; Search time 46 Seconds  
(without alignments)  
3747.327 Million cell updates/sec

Title: US-09-720-383C-10  
Perfect score: 5918  
Sequence: 1 MEASAGLVAGSHNRNRLVLI.....DPFLAKDGPDLLECGLCDN 1086

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq 19Jun03:\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5918	100.0	1086	21	AAV84111 Amino acid sequenc
2	5918	100.0	1086	21	AAV84116 Amino acid sequenc
3	5918	100.0	1086	21	AAV58835 Corn cellulose syn
4	5570.5	94.1	1148	21	AAV58832 Corn cellulose syn
5	5402.5	91.3	1058	21	AAV84107 Amino acid sequenc
6	5228	88.3	1094	21	AAV84113 Amino acid sequenc
7	5228	88.3	1094	21	AAV84118 Amino acid sequenc
8	5228	88.3	1165	21	AAV58834 Corn cellulose syn
9	4278.5	72.3	1084	23	ABB93304 Herbicidally activ

10	4262	72.0	1084	23	ABB93949 Herbicidally activ
11	4250.5	71.8	1084	19	AAW33818 Arabidopsis cellul
12	4209.5	71.1	1088	23	ABB91831 Herbicidally activ
13	4204.5	71.0	1069	23	ABB93411 Herbicidally activ
14	4185.5	70.7	1080	22	AAW37895 Arabidopsis thalia
15	3986	67.4	1039	21	AAV58837 Soybean cellulose
16	3984	67.3	1075	21	AAV84109 Amino acid sequenc
17	3984	67.3	1075	21	AAV84115 Amino acid sequenc
18	3984	67.3	1075	21	AAV84120 Amino acid sequenc
19	3977.5	67.2	1074	21	AAV84110 Amino acid sequenc
20	3977.5	67.2	1074	21	AAV84112 Amino acid sequenc
21	3977.5	67.2	1074	21	AAV84117 Amino acid sequenc
22	3967.5	67.0	1026	23	ABB93522 Arabidopsis cellul
23	3945.5	66.7	1081	19	AAW33817 Arabidopsis cellul
24	3945.5	66.7	1081	23	ABB93199 Arabidopsis cellul
25	3935.5	66.5	1081	19	AAW33820 Arabidopsis cellul
26	3911	66.1	1076	21	AAV84121 DNA encoding a mai
27	3903.5	66.0	1039	19	AAW73309 Cellulose synthase
28	3899.5	65.9	1077	21	AAV84108 Amino acid sequenc
29	3899.5	65.9	1077	21	AAV84114 Amino acid sequenc
30	3898.5	65.9	1065	21	AAW48733 Arabidopsis thalia
31	3898.5	65.9	1065	23	ABB93356 Arabidopsis thalia
32	3887.5	65.7	1054	21	AAW48734 Arabidopsis thalia
33	3880.5	65.6	1079	23	AAU10496 Corn cellulose syn
34	3876.5	65.5	1065	19	AAW33819 Arabidopsis cellul
35	3804.5	64.3	1043	23	ABB93701 Arabidopsis cellul
36	3782	63.9	1065	23	ABB91880 Arabidopsis cellul
37	3780.5	63.9	1043	21	AAV84119 Arabidopsis cellul
38	3603	60.9	959	21	AAW48735 Arabidopsis cellul
39	3521	59.5	793	21	AAV58839 Arabidopsis thalia
40	3476.5	58.7	974	19	AAW73308 Soybean cellulose
41	3469.5	58.6	974	19	AAW60039 Cotton cellulose s
42	3356	56.7	955	21	AAW16338 Eucalyptus grandis
43	3329.5	56.3	958	23	ABB92981 Herbicidally activ
44	3318	56.1	821	23	AAU10495 Corn cellulose syn
45	3138.5	53.0	946	22	AAW37893 Populus tremuloide

ALIGNMENTS

RESULT 1  
ID AAV84111 standard; Protein; 1086 AA.  
XX AAV84111;  
AC  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DB Amino acid sequence of a maize cellulose synthase.  
XX  
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker.  
XX  
OS Zea mays.  
XX  
PN WO200009706-A2.  
XX  
PD 24-FEB-2000.  
XX  
XX  
PF 16-AUG-1999; 99WO-US18760.  
XX  
PR 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
PI Dhuga KS, Helentjaris TG, Bowen BA, Wang X;  
XX WPI; 2000-224343/19.  
XX N-PSDB; AA299503.  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants



PT and to produce transgenic plants expressing the novel protein -  
XX  
PS Claim 15; Page 118-120; 119pp; English.  
XX  
CC The present sequence represents a maize cellulose synthase polypeptide.  
CC The cellulose synthase can be used for the improvement of stalk quality  
CC for improved stand or silage. It also provides an increased concentration  
CC of cellulose in the pericarp, hardening the kernel and improving its  
CC handling ability. The sequences are used to produce transgenic plants  
CC and seeds expressing the cellulose synthase. The polynucleotide is  
CC used for modulating, preferably increasing, the level of the synthase  
CC in a plant cell. The plants are preferably monocots. The polynucleotide  
CC is also used as a probe or primer in the detection quantitation or  
CC isolation of gene transcripts. The probes are useful in detecting  
CC deficiencies in the level of mRNA in screenings for desired transgenic  
CC plant, for detecting mutations in the gene, for monitoring upregulation  
CC of expression or changes in enzyme activity in screening assays of  
CC compounds, for detection of any number of allelic variants of the gene,  
CC or for use as molecular markers in plant breeding programs. The  
CC isolated nucleic acids of the present invention can also be used for  
CC recombinant expression of their encoded polypeptides or for use as  
CC immunogens in the preparation and/or screening of antibodies. The  
CC proteins can be employed in assays for enzyme agonists or antagonists  
CC of enzyme function or for use of immunogens or antigens to obtain  
CC antibodies specifically immunoreactive with a protein.  
XX  
SQ Sequence 1086 AA;  
Query Match 100.0%; Score 5918; DB 21; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREONGVCQICGDDVGLAPGDDPFVACNE 60  
DB 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREONGVCQICGDDVGLAPGDDPFVACNE 60  
QY 61 CAFVPCRDCEYERREGTONCPQCKTRYKRLKGCQRTVGDEEEDGVDDLNEFNMDGHS 120  
DB 61 CAFVPCRDCEYERREGTONCPQCKTRYKRLKGCQRTVGDEEEDGVDDLNEFNMDGHS 120  
QY 121 QSVAESMLYGHMSYGRGDDPNGAPOAFQLNPNVPLLTNGQVDDIPPEQHALVPSFMGG 180  
DB 121 QSVAESMLYGHMSYGRGDDPNGAPOAFQLNPNVPLLTNGQVDDIPPEQHALVPSFMGG 180  
QY 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAWKERMENWKQDERMHTGNDGG 240  
DB 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAWKERMENWKQDERMHTGNDGG 240  
QY 241 DDGDDADLPLMDEARQOLSRIKPLPSSQINPYRMIIIRLVLGFFFHYRVMHPVNDABA 300  
DB 241 DDGDDADLPLMDEARQOLSRIKPLPSSQINPYRMIIIRLVLGFFFHYRVMHPVNDABA 300  
QY 301 LMLISVICEIWFAMSMILDOFPKWFPIERETVLDRLSLRFDKEGQPSQLAIDFVSTVD 360  
DB 301 LMLISVICEIWFAMSMILDOFPKWFPIERETVLDRLSLRFDKEGQPSQLAIDFVSTVD 360  
QY 361 PLKEPPLVTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSEPAKKWVPFCRY 420  
DB 361 PLKEPPLVTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSEPAKKWVPFCRY 420  
QY 421 NIEPRAPEWYFOQKIDYLDKDVAAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480  
DB 421 NIEPRAPEWYFOQKIDYLDKDVAAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480  
QY 481 MODGTPWPGNNVBDHGMIOVFLGSGGLDCEGNEPRLVYVSREKRPQYHHKKAGAMN 540  
DB 481 MODGTPWPGNNVBDHGMIOVFLGSGGLDCEGNEPRLVYVSREKRPQYHHKKAGAMN 540  
QY 541 ALVRSASVLTNAPYLLNDGDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600  
DB 541 ALVRSASVLTNAPYLLNDGDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600  
QY 601 HDRYANRNVVFIDINMKGLDGIQGPITYGTGCVFRRQALYGYDAPKTKPPSRTCNCMPK 660

DB 601 HDRYANRNVVFIDINMKGLDGIQGPITYGTGCVFRRQALYGYDAPKTKPPSRTCNCMPK 660  
QY 661 WCFCCCFGNRKQKTKTKTEKKLLFPKKEENQSPAYALGEIDEAPGAENKAGIYN 720  
DB 661 WCFCCCFGNRKQKTKTKTEKKLLFPKKEENQSPAYALGEIDEAPGAENKAGIYN 720  
QY 721 QOKLEKFGQSSVFTSTLENGSTLKSASPASLKEAIIHVISCGYEDKTWGEIGWY 780  
DB 721 QOKLEKFGQSSVFTSTLENGSTLKSASPASLKEAIIHVISCGYEDKTWGEIGWY 780  
QY 781 GSVTEBILTGFKMHCHGWRSTIYCI PKRVAFKGSAPLNSDRLHQVLRWALGSEIFPSNH 840  
DB 781 GSVTEBILTGFKMHCHGWRSTIYCI PKRVAFKGSAPLNSDRLHQVLRWALGSEIFPSNH 840  
QY 841 CPLMYGYGGGLKFLERFSYINSIVYPWTSIPLAYCTIPAICLLTGKFTPELNVA SLW 900  
DB 841 CPLMYGYGGGLKFLERFSYINSIVYPWTSIPLAYCTIPAICLLTGKFTPELNVA SLW 900  
QY 901 FMSLFICIFATSILEMWGSVIGIDWWRNEQFWVIGVSHLFAVFOGLKVIAGVDTSF 960  
DB 901 FMSLFICIFATSILEMWGSVIGIDWWRNEQFWVIGVSHLFAVFOGLKVIAGVDTSF 960  
QY 961 TVTSKGDDEEFSELYTFKWTLLIPTTLLLNFTIGVAVGSNAINNGESWGPFGKL 1020  
DB 961 TVTSKGDDEEFSELYTFKWTLLIPTTLLLNFTIGVAVGSNAINNGESWGPFGKL 1020  
QY 1021 FFAFWYIVHLYPFLKGLVGRQNRPTIVIVWSILASIFSLMWRIDPFLAKDDGPLLEE 1080  
DB 1021 FFAFWYIVHLYPFLKGLVGRQNRPTIVIVWSILASIFSLMWRIDPFLAKDDGPLLEE 1080  
QY 1081 CGLDCN 1086  
DB 1081 CGLDCN 1086  
RESULT 2  
AA84116  
ID AA84116 standard; Protein; 1086 AA.  
XX  
AC AA84116;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE Amino acid sequence of a maize cellulose synthase.  
XX  
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker.  
XX  
OS Zea mays.  
XX  
PN WO200009706-A2.  
XX  
PD 24-FEB-2000.  
XX  
PF 16-AUG-1999; 99WO-US18760.  
XX  
PR 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Dhurga KS, Helentjaris TG, Bowen BA, Wang X;  
XX  
DR WPI; 2000-224343/19.  
DR N-PSDB; AA299518.  
XX  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein -  
XX  
PS Claim 15; Page 157-160; 119pp; English.  
CC The present sequence represents a maize cellulose synthase polypeptide.

CC The cellulose synthase can be used for the improvement of stalk quality  
CC for improved stand or silage. It also provides an increased concentration  
CC of cellulose in the pericarp, hardening the kernel and improving its  
CC handling ability. The sequences are used to produce transgenic plants  
CC and seeds expressing the cellulose synthase. The polynucleotide is  
CC used for modulating, preferably increasing, the level of the synthase  
CC in a plant cell. The plants are preferably monocots. The polynucleotide  
CC is also used as a probe or primer in the detection quantitation or  
CC isolation of gene transcripts. The probes are useful in detecting  
CC deficiencies in the level of mRNA in screenings for desired transgenic  
CC plant, for detecting mutations in the gene, for monitoring upregulation  
CC of expression or changes in enzyme activity in screening assays of  
CC compounds, for detection of any number of allelic variants of the gene,  
CC or for use as molecular markers in plant breeding programs. The  
CC isolated nucleic acids of the present invention can also be used for  
CC recombinant expression of their encoded polypeptides or for use as  
CC immunogens in the preparation and/or screening of antibodies. The  
CC proteins can be employed in assays for enzyme agonists or antagonists  
CC of enzyme function or for use of immunogens or antigens to obtain  
CC antibodies specifically immunoreactive with a protein.

SO Sequence 1086 AA;

Query Match 100.0%; Score 5918; DB 21; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEASAGLVAGSHNRNELVIRDDGEPKPPREQNGVCQICGDDVGLAPGDPFVACNE 60  
Db 1 MEASAGLVAGSHNRNELVIRDDGEPKPPREQNGVCQICGDDVGLAPGDPFVACNE 60  
QY 61 CAFVYCRDCYEYERREGTONCPCKTRYKRLKGCQRVTGDEEDGVDDLDFNWDGHS 120  
Db 61 CAFVYCRDCYEYERREGTONCPCKTRYKRLKGCQRVTGDEEDGVDDLDFNWDGHS 120  
QY 121 QSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQWVDIIPRQHALVPSFMGG 180  
Db 121 QSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQWVDIIPRQHALVPSFMGG 180  
QY 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMENTWKORERMHOTGNDGG 240  
Db 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMENTWKORERMHOTGNDGG 240  
QY 241 DGDGDADLPMDPEARQOLSRKIPRSSQINPYRMIIRLVVLGFFHYRVMHPVNDFA 300  
Db 241 DGDGDADLPMDPEARQOLSRKIPRSSQINPYRMIIRLVVLGFFHYRVMHPVNDFA 300  
QY 301 LWLISVCEIWFAMSMWILDQFPKPIERETYLDRLSLRFDEGQPSQLAPIDFVSTVD 360  
Db 301 LWLISVCEIWFAMSMWILDQFPKPIERETYLDRLSLRFDEGQPSQLAPIDFVSTVD 360  
QY 361 PLKEPPLVTNTVLSILSVDPVDVKVSCYVSDDGAAMLTFEALSETSEPAKMWPFCKRY 420  
Db 361 PLKEPPLVTNTVLSILSVDPVDVKVSCYVSDDGAAMLTFEALSETSEPAKMWPFCKRY 420  
QY 421 NIEPRAPEWYFOQKIDYLDKQVAAVFERRAMKREYEEFKVRINALVAKAQVPEEGWT 480  
Db 421 NIEPRAPEWYFOQKIDYLDKQVAAVFERRAMKREYEEFKVRINALVAKAQVPEEGWT 480  
QY 481 MODGTPWPGNNVRDHPGMIQVFLQSGGLDCEGNELPRLVYVSREKRPGYNHKKAGAMN 540  
Db 481 MODGTPWPGNNVRDHPGMIQVFLQSGGLDCEGNELPRLVYVSREKRPGYNHKKAGAMN 540  
QY 541 ALVRYSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600  
Db 541 ALVRYSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600  
QY 601 HDRYANRVVFPDINMGLDGIQGPVYGTGCVFRQALYGYDAPKTKPPSRTCNCWPK 660  
Db 601 HDRYANRVVFPDINMGLDGIQGPVYGTGCVFRQALYGYDAPKTKPPSRTCNCWPK 660  
QY 661 WCFCCCCFGRNRKQKTKTKPTEKKLLFFKKEENQSPAYALGEIDEAAPGAENKAGIVN 720  
Db 661 WCFCCCCFGRNRKQKTKTKPTEKKLLFFKKEENQSPAYALGEIDEAAPGAENKAGIVN 720

Db 661 WCFCCCCFGRNRKQKTKTKPTEKKLLFFKKEENQSPAYALGEIDEAAPGAENKAGIVN 720  
QY 721 QQLEKKFGQSSVFVSTLLENGTLLKASAPASLKEAIIHVISCGYEDKTDWKEIGWY 780  
Db 721 QQLEKKFGQSSVFVSTLLENGTLLKASAPASLKEAIIHVISCGYEDKTDWKEIGWY 780  
QY 781 GSVTEDILTGFKMHCHGMRISYICPKRVAFKGSAPNLSDRLHQVLRWALGSIIFFSNH 840  
Db 781 GSVTEDILTGFKMHCHGMRISYICPKRVAFKGSAPNLSDRLHQVLRWALGSIIFFSNH 840  
QY 841 CPLWYGYGGGLKFLERSYINSIVYPWTSIPLAYCTLLPAICLLTGKFTPBELNVASLW 900  
Db 841 CPLWYGYGGGLKFLERSYINSIVYPWTSIPLAYCTLLPAICLLTGKFTPBELNVASLW 900  
QY 901 FMSLFICIFATSILEMRSVGVGIDWWRNEQFWYIGVSSHLPFAVFOGLLKVIAGVDTSF 960  
Db 901 FMSLFICIFATSILEMRSVGVGIDWWRNEQFWYIGVSSHLPFAVFOGLLKVIAGVDTSF 960  
QY 961 TVTSKGGDDEEFSELYTFKWTLLIPTTLLLNFIQVAVASNAIINNGYESWGPLFGKL 1020  
Db 961 TVTSKGGDDEEFSELYTFKWTLLIPTTLLLNFIQVAVASNAIINNGYESWGPLFGKL 1020  
QY 1021 FFAFWIVHLYPFLKGLVGRQNRTPTIIVWSILLASIFSLMWRIDPFLAKDGPILLE 1080  
Db 1021 FFAFWIVHLYPFLKGLVGRQNRTPTIIVWSILLASIFSLMWRIDPFLAKDGPILLE 1080  
QY 1081 CGLDCN 1086  
Db 1081 CGLDCN 1086

RESULT 3

AAY58835  
ID AAY58835 standard; Protein; 1086 AA.

AC AAY58835;

DT 08-MAY-2000 (first entry)

DE Corn cellulose synthase (full-length).

KM Corn; maize; cellulose synthase; transgenic plant.

OS Zea mays.

PN WO200004166-A2.

PD 27-JAN-2000.

PF 13-JUL-1999; 99WO-US15871.

PR 14-JUL-1998; 98US-0092844.

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;  
PI Rafalski JA, Thorpe CJ;

XX MPI; 2000-182431/16.

DR N-PSDB; AA258266.

PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
PT probes for isolating cDNAs and genes encoding homologous proteins, for  
PT producing transgenic plants -

XX Claim 12; Page 50-53; 93pp; English.

PS The present sequence is that of full-length corn cellulose

CC synthase (CS), as deduced from an isolated cDNA clone (see AAY58266).

CC The invention relates to isolated CS nucleic acid fragments and to

CC CS polypeptides. It also relates to the construction of a chimeric

CC gene encoding all or a portion of the CS, in sense or antisense

CC orientation, where expression of the gene results in altered levels

CC of the CS in transformed host cells. The host cells can be used to  
CC screen compounds for their ability to inhibit CS activity. CS  
CC nucleic acids are also useful for producing transgenic plants having  
CC altered levels of CS, and hence altered levels of fibre. CS may  
CC also serve as a target for the development of novel herbicides.

XX  
SQ Sequence 1086 AA;

Query Match 100.0%; Score 5918; DB 21; Length 1086;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCQICGDDVGLAPGDPFVACNE 60  
DB 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCQICGDDVGLAPGDPFVACNE 60  
QY 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEEDGVDDLNEFNMDGHS 120  
DB 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEEDGVDDLNEFNMDGHS 120  
QY 121 QSVAESMLYGHMSYGRGDDPNGAPOAFQLNPNVPLLTNGQMVDDIPPEQHALLVPSFMGG 180  
DB 121 QSVAESMLYGHMSYGRGDDPNGAPOAFQLNPNVPLLTNGQMVDDIPPEQHALLVPSFMGG 180  
QY 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAWKERMENWKQRQERHQTGNDGG 240  
DB 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAWKERMENWKQRQERHQTGNDGG 240  
QY 241 DDGDDADLPLMDEARQOLSRIPLPSSQINPYRMIIRLVVLGFFHYRVMPVNDADA 300  
DB 241 DDGDDADLPLMDEARQOLSRIPLPSSQINPYRMIIRLVVLGFFHYRVMPVNDADA 300  
QY 301 LMLISVICEIWFAMSWILDQFPKWFPIERETYLRLSLRFDKEGQPSQLAPIDFVSTVD 360  
DB 301 LMLISVICEIWFAMSWILDQFPKWFPIERETYLRLSLRFDKEGQPSQLAPIDFVSTVD 360  
QY 361 PLKEPPLVTNTVLSILSVDPVDKVSQVSDGAMLTFEALSETSEFAKKMVPFCRY 420  
DB 361 PLKEPPLVTNTVLSILSVDPVDKVSQVSDGAMLTFEALSETSEFAKKMVPFCRY 420  
QY 421 NIEPRAPWYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALVAKQVPEEGWT 480  
DB 421 NIEPRAPWYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALVAKQVPEEGWT 480  
QY 481 MODGTPWPGNNVRDHGMIQVFLGSGGLDCEGNELPRLVYVSREKRPGYNHKKAGAMN 540  
DB 481 MODGTPWPGNNVRDHGMIQVFLGSGGLDCEGNELPRLVYVSREKRPGYNHKKAGAMN 540  
QY 541 ALRVSAVLTNAPVLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQFQRFIDIDR 600  
DB 541 ALRVSAVLTNAPVLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQFQRFIDIDR 600  
QY 601 HDRYANRNVFEDINMKGLDGIQGPITYGTGCFRRQALYGYDAPKTKKPPSRTCNCMPK 660  
DB 601 HDRYANRNVFEDINMKGLDGIQGPITYGTGCFRRQALYGYDAPKTKKPPSRTCNCMPK 660  
QY 661 WCFCCCCFGRNRKQKTKTKTEKKLLFFKKEENSPAYALGEIDEAAPGAENEKAGIVN 720  
DB 661 WCFCCCCFGRNRKQKTKTKTEKKLLFFKKEENSPAYALGEIDEAAPGAENEKAGIVN 720  
QY 721 QOKLEKKFGQSSVFVTSSTLLENGGTLKSASPASLKEAIIHVISCGYEDKTDWKEIGMIY 780  
DB 721 QOKLEKKFGQSSVFVTSSTLLENGGTLKSASPASLKEAIIHVISCGYEDKTDWKEIGMIY 780  
QY 781 GSVTEDILTGFKMHCHGWSIYCIPIKRVAFKGSAPLNLSDRLHQVLRWALGSIIEFFSNH 840  
DB 781 GSVTEDILTGFKMHCHGWSIYCIPIKRVAFKGSAPLNLSDRLHQVLRWALGSIIEFFSNH 840  
QY 841 CPLWYGYGGGLKFLERFSYINSIVPMTSIFPLAYCTLLPAICLLTGKFTPELNNVASLW 900  
DB 841 CPLWYGYGGGLKFLERFSYINSIVPMTSIFPLAYCTLLPAICLLTGKFTPELNNVASLW 900  
QY 901 FMSLFICIFATSILEMRSQGVGIDDDWMRNEQFWVIGVSSHLFAVFOGLKVIAGVDTSF 960

DB 901 FMSLFICIFATSILEMRSQGVGIDDDWMRNEQFWVIGVSSHLFAVFOGLKVIAGVDTSF 960  
QY 961 TVTSKGGDDEEFSSELYTFKWTLLIPPTLLLNFIQVAVGSNAINNGYESWGPFGKL 1020  
DB 961 TVTSKGGDDEEFSSELYTFKWTLLIPPTLLLNFIQVAVGSNAINNGYESWGPFGKL 1020  
QY 1021 FFAFWIVHLYPFLKGLVGRNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDGPLEE 1080  
DB 1021 FFAFWIVHLYPFLKGLVGRNRTPTIVIVWSILLASIFSLLWVRIDPFLAKDGPLEE 1080  
QY 1081 CGLDCN 1086  
DB 1081 CGLDCN 1086

RESULT 4  
AA558832  
ID AA558832 standard; Protein; 1148 AA.

AC AA558832;  
XX  
DT 08-MAY-2000 (first entry)  
XX  
DE Corn cellulose synthase (partial sequence).  
XX  
KM Corn; maize; cellulose synthase; transgenic plant.  
XX  
OS Zea mays.  
XX  
PN WO200004166-A2.  
XX  
PD 27-JAN-2000.  
XX  
PF 13-JUL-1999; 99WO-US15871.  
XX  
PR 14-JUL-1998; 98US-0092844.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;  
PI Rafalski JA, Thorpe CJ;  
XX  
DR WPI; 2000-182431/16.  
DR N-PSDB; AA258263.  
XX

PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
PT probes for isolating cDNAs and genes encoding homologous proteins, for  
PT producing transgenic plants -  
XX  
PS Claim 12; Page 39-42; 93pp; English.  
XX

CC The present sequence is that of a portion of corn cellulose  
CC synthase (CS), as deduced from a contig of cDNA clones (see AA558263).  
CC The invention relates to isolated CS nucleic acid fragments and to  
CC CS polypeptides. It also relates to the construction of a chimeric  
CC gene encoding all or a portion of the CS, in sense or antisense  
CC orientation, where expression of the gene results in altered levels  
CC of the CS in transformed host cells. The host cells can be used to  
CC screen compounds for their ability to inhibit CS activity. CS  
CC nucleic acids are also useful for producing transgenic plants having  
CC altered levels of CS, and hence altered levels of fibre. CS may  
CC also serve as a target for the development of novel herbicides.

XX  
SQ Sequence 1148 AA;

Query Match 94.1%; Score 5570.5; DB 21; Length 1148;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 1027; Conservative 23; Mismatches 35; Indels 5; Gaps 3;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCQICGDDVGLAPGDPFVACNE 60  
DB 60 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVCQICGDDVGRNPDGEFVACNE 119



QY 61 CAFPVCRDCYEREREGTQNCPOCKTRYKLGCCQRTGTDEEDGVDDLNEFNW-DGHD 119  
DB 120 CAFPICRDCYEREREGTQNCPOCKTRFKLKGCAVPGDEEDGVDDLNEFNWSDKHD 179  
QY 120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQOLNPNVPLLTNGQMVDDIPPEQHALVPSFMG 179  
DB 180 SQYLAESMLHAHMSYGRGADLDGVPQPFHPINVPPLLTNGQMVDDIPPDQHALVPSFVG 239  
QY 180 GGRIRHPLPYADPSLPVQPRSMPSKDLAAYGYSVAMKERMENWKQORMHQTGNDG 239  
DB 240 GGRIRHPLPYADPNLPVQPRSMPSKDLAAYGYSVAMKERMESWKQORMHQTGNDG 299  
QY 240 GDDGDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMHPVNDAF 299  
DB 300 GDDGDADPLMDEARQPLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMHPVDAF 359  
QY 300 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDEKQPSQLAPIDFVSTV 359  
DB 360 ALWLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDEKQPSQLAPVDFVSTV 419  
QY 360 DPLKEPPLVTNTVLSLSVDYPVDKVCYVSDGAAMLTPEALSTSEFAKKWVPFCKR 419  
DB 420 DPLKEPPLVTANTVLSLSVDYPVDKVCYVSDGAAMLTPEALSTSEFAKKWVPFCKR 479  
QY 420 YNIEPRAPEWYFOQKIDYLDKQVAANFVRERRAMKREYEEFKVRINALVAKQKPEEGW 479  
DB 480 YSLEPRAPEWYFOQKIDYLDKQVAPNFRERRAMKREYEEFKVRINALVAKQKPEEGW 539  
QY 480 TMDGTWPWGNVNRDHPGMIQVFLGSGGLDCEGNELPRLVYVSREKRGYNHKKAGAM 539  
DB 540 TMDGTWPWGNVNRDHPGMIQVFLGSGGHVGENELPRLVYVSREKRGYNHKKAGAM 599  
QY 540 NALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMDPLGKKVCYVQFPQRFDGID 599  
DB 600 NALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMDPLGKKVCYVQFPQRFDGID 659  
QY 600 RHDRYANRNVVFFDINMKGLDGIQPIYVGTGCVFRROALYGYDAPTKKPPSRTCNCWP 659  
DB 660 RHDRYANRNVVFFDINMKGLDGIQPIYVGTGCVFRROALYGYDAPTKKPPSRTCNCWP 719  
QY 660 KWCFCFCCCFEGNKKQ--KTKPKTEKKLLFFKKEENQSPAYALGEIDEAAPGAENKA 716  
DB 720 KWCICCCCFGNRKTKKTKTSKPKFEKIKKI--FKKENQAPAYALGEIDEAAPGAENKA 778  
QY 717 GIWNOQKLEKKFGSSVFTSTLLENGTILKSASPASLLKEAIIHVISGGEYEDKTDWKEI 776  
DB 779 SIWNOQKLEKKFGSSVFASTLLENGTILKSASPASLLKEAIIHVISGGEYEDKTDWGXDI 838  
QY 777 GMIYGSVTEDLITGFKMHCHGWSIYCIPIKRAAFKGSAPLNLSDRLHOVLKRWALGSIEIF 836  
DB 839 GMIYGSVTEDLITGFKMHCHGWSIYCIPIKRAAFKGSAPLNLSDRLHOVLKRWALGSIEIF 898  
QY 837 FSNHCPLMWYGGGLKFLERFSYINSIVPWTISIPLAYCTLPALCILLTGKFITPELANV 896  
DB 899 FSNHCPLMWYGGGLKFLERFSYINSIVPWTISIPLAYCTLPALCILLTGKFITPELANV 958  
QY 897 ASLWFMSLFCIFATSILEMRWSGVGIDMWRENFQFVIGVSSHFAVFOGLLKVIAGV 956  
DB 959 ASLWFMSLFCIFATSILEMRWSGVGIDMWRENFQFVIGVSSHFAVFOGLLKVIAGV 1018  
QY 957 DTSFTVTSKGGDDEEFSELYTFKWTLLIPPTLLLNFIQVAVGASNAINNGYESWGPL 1016  
DB 1019 DTSFTVTSKGGDDEEFSELYTFKWTLLIPPTLLLNFIQVAVGASNAINNGYESWGPL 1078  
QY 1017 FGKLFFAFWIVHLYPFLKGLVGRQNRPTIIVWSILLASISFLWVRIDPFLAKDDGP 1076  
DB 1079 FGKLFFAFWIVHLYPFLKGLVGRQNRPTIIVWSILLASISFLWVRIDPFLAKDDGP 1138  
QY 1077 LLEBCGLDCN 1086  
DB 1139 LLEBCGLDCN 1148

RESULT 5  
AAAY84107  
ID AAY84107 standard; Protein; 1058 AA.  
XX  
AC AAY84107;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE Amino acid sequence of a maize cellulose synthase.  
XX  
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker.  
XX  
OS Zea mays.  
XX  
PN WO200009706-A2.  
XX  
PD 24-FEB-2000.  
XX  
PF 16-AUG-1999; 99WO-US18760.  
XX  
PR 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
XX  
DR WPI; 2000-224343/19.  
DR N-PSDB; AA299491.  
XX  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants  
XX  
PS Claim 15; Page 87-89; 119pp; English.  
XX  
CC The present sequence represents a maize cellulose synthase polypeptide.  
CC The cellulose synthase can be used for the improvement of stalk quality  
CC for improved stand or silage. It also provides an increased concentration  
CC of cellulose in the pericarp, hardening the kernel and improving its  
CC handling ability. The sequences are used to produce transgenic plants  
CC and seeds expressing the cellulose synthase. The polynucleotide is  
CC used for modulating, preferably increasing, the level of the synthase  
CC in a plant cell. The plants are preferably monocots. The polynucleotide  
CC is also used as a probe or primer in the detection quantitation or  
CC isolation of gene transcripts. The probes are useful in detecting  
CC deficiencies in the level of mRNA in screenings for desired transgenic  
CC plant, for detecting mutations in the gene, for monitoring upregulation  
CC of expression or changes in enzyme activity in screening assays of  
CC compounds, for detection of any number of allelic variants of the gene,  
CC or for use as molecular markers in plant breeding programs. The  
CC isolated nucleic acids of the present invention can also be used for  
CC recombinant expression of their encoded polypeptides or for use as  
CC immunogens in the preparation and/or screening of antibodies. The  
CC proteins can be employed in assays for enzyme agonists or antagonists  
CC of enzyme function or for use of immunogens or antigens to obtain  
CC antibodies specifically immunoreactive with a protein.  
XX  
SQ Sequence 1058 AA;  
  
Query Match 91.3%; Score 5402.5; DB 21; length 1058;  
Best Local Similarity 94.0%; Pred. No. 0;  
Matches 994; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

QY 33 EQNGVQICGDDVGLAPGDPFVACNECAPVCRDCYEREREGTQNCPOCKTRYKLG 92  
DB 3 QRNQVQICGDDVGRNPDGDPFVACNECAPICRDCYEREREGTQNCPOCKTRFK 62  
QY 93 GCQVYTGDEEDGVDDLNEFNW-DGHDQSVAESMLYGHMSYGRGGDPNGAPQAFQLN 151  
DB 63 GCARVPGDEEDGVDDLNEFNWSDKHDSQYLAESMLHAHMSYGRGADLDGVPQPFHP 122

QY	152	NVPLLTNGQWDDIPPEQHALLVPSFMGGGCKRIHPLPYADPSLPVQPRSMDBSKDLAAYG	211
Db	123	NVPLLTNGQWDDIPPDQHALVPSFVGCGGCKRIHPLPYADPNLPVQPRSMDBSKDLAAYG	182
QY	212	YGSVAMKERMENWKOBERMHQTGNDGGDDGDADPLMDEARQOLSRKIPLBSSQINP	271
Db	183	YGSVAMKERMESWKQKERMHQTRNDGGDDGDADPLMDEARQPLSRKIPLBSSQINP	242
QY	272	YRMIIIRLVVLGFFHHYRVMHPVNDAFALWLISVCEIWFAMSWILDOFPKMFPIERET	331
Db	243	YRMIIIRLVVLGFFHHYRVMHPVPDAFALWLISVCEIWFAMSWILDOFPKMFPIERET	302
QY	332	YLDRLSLRFDKEGQPSQLAPIDFFVSTVDPLKEPPLVTNTVL SILSVDPVDKVS CYVS	391
Db	303	YLDRLSLRFDKEGHPSQLAPVDFVSTVDPLKEPPLVTANTVLSILSVDPVDKVS CYVS	362
QY	392	DDGAAMLTFEALSETSEFAKKWVPFCRKYNIERAPENYFQOKIDYLDKQVAANFVRER	451
Db	363	DDGAAMLTFEALSETSEFAKKWVPFCRKYSLERAPENYFQOKIDYLDKQVAANFVRER	422
QY	452	AMKREYEEFKVRINALVAKAQKVPBEGWTMDGTPMGNVNRDHPGMIQVFLGSGGLDC	511
Db	423	AMKREYEEFKVRINALVAKAQKVPBEGWTMDGTPMGNVNRDHPGMIQVFLGSGGHDV	482
QY	512	EGNELPRLVYVSREKRPNGNHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIK	571
Db	483	EGNELPRLVYVSREKRPNGNHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIK	542
QY	572	EAMCFMMDPLGKKVCYVQPPQRFDGIDRHDYANRNVYFEDINMGLDGIQGPVYGTG	631
Db	543	EAMCFMMDPLGKKVCYVQPPQRFDGIDRHDYANRNVYFEDINMGLDGIQGPVYGTG	602
QY	632	CVFRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCCFGNRKOK--KTTKPKTEKKLLF	688
Db	603	CVFRQALYGYDAPKTKKPPSRTCNCWPKWCICCCCFGNRKTKKTKTSKPKFEKIKKL-	661
QY	689	EKKEENQSPAYALGEIDEAAGAENEKAGIVNOQKLEKFGQSSVFTSTLLENGTLKS	748
Db	662	EKKEENQAPAYALGEIDEAAGAENEKASIVNOQKLEKFGQSSVFASTLLENGTLKS	721
QY	749	ASPASLLKEAIIHVISCGYEDKTDWGKEIGMIYGSVTEBILTGFKMHCHGWSIYCIPIKRV	808
Db	722	ASPASLLKEAIIHVISCGYEDKTDWGKDIGMIYGSVTEBILTGFKMHCHGWSIYCIPIKRA	781
QY	809	AFKGSAPLNLSDRLHQVLRWALGSIIEIFSNHCPLMWYGGGLKFLERFSYINSIYVPMT	868
Db	782	AFKGSAPLNLSDRFHQVLRWALGSIIEIFSNHCPLMWYGGGLKFLERFSYINSIYVPMT	841
QY	869	SIPPLAYCTLPALICLLTGKFTPELNNVASLWMFMSLFCIPATSILEMWMSGVGIDDMWR	928
Db	842	SIPPLAYCTLPALICLLTGKFTPELNNVASLWMFMSLFCIPATSILEMWMSGVGIDDMWR	901
QY	929	NEQFWIVIGVSSHLPFAVFOGLLKVIAGVDTSFTVTSKGDDEEFSELYTFKMTTLLIPT	988
Db	902	NEQFWIVIGVSSHLPFAVFOGLLKVIAGVDTSFTVTSKGDDEEFSELYTFKMTTLLIPT	961
QY	989	TLLLLNFIGVAVGSNAINNGYESWGPLFGKLFPAFVIVHLXPFLKGLVGRQNRPTIV	1048
Db	962	TLLLLNFIGVAVGISNAINNGYESWGPLFGKLFPAFVIVHLXPFLKGLVGRQNRPTIV	1021
QY	1049	IWMSILLASIFSLMWRIDPFLAKDGPPLBECGLDC	1085
Db	1022	IWMSILLASIFSLMWRIDPFLAKDGPPLBECGLDC	1058

RESULT 6	
AAV84113	
ID	AAV84113 standard; Protein; 1094 AA.
XX	
AC	AAV84113;
XX	
DT	03-JUL-2000 (first entry)
XX	

DE Amino acid sequence of a maize cellulose synthase.  
XX  
XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker.  
KW  
XX  
OS Zea mays.  
XX  
XX WO200009706-A2.  
PN  
XX  
PD 24-FEB-2000.  
XX  
XX 16-AUG-1999; 99WO-US18760.  
PF  
XX  
XX 17-AUG-1998; 98US-0096822.  
PR  
XX  
XX (PION-) PIONEER HI-BRED INT INC.  
PA  
XX  
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
PI  
XX  
XX WPI; 2000-224343/19.  
DR N-PSDB; AAZ99509.  
DR  
XX  
XX  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein -  
XX  
XX  
PS Claim 15; Page 134-136; 119pp; English.

CC The present sequence represents a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain  
 CC antibodies specifically immunoreactive with a protein.

	Query Match	88.3%;	Score 5228;	DB 21;	Length 1094;	
	Best Local Similarity	88.7%;	Pred. No. 0;			
	Matches 974;	Conservative 38;	Mismatches 70;	Indels 16;	Gaps 7	
QY	1 MEASAGLVAGSHNRNELVIRRD--GDPGPKPREONGVCQICGDDVGLAPGGDFVA	57				
Dd	1 MEASAGLVAGSHNRNELVIRDRRESGAAGGARRAEAP-CQICGEVGVFDEEPFVA	59				
QY	58 CNECAFVPCRCDCYEYERREGTONCPQCCKTRYKRLKGCCQRTGTDEEDDGVDDLNEFN-WD	116				
Dd	60 CNECAFPVCRACCYEYERREGSQACPQCCKTRYKRLKGCPRVAGDEEDDGVDDLEGFQLD	119				
QY	117 G----HDSQSVAESMLYGHMSYGRGDPNGAQAFAQLNPNPPLLTNGQMVDIIPPEQHAI	172				
Dd	120 GAAHEDDPQYVAESMLRAQMSYGRGD---AHFGFSVPNPVPLLTNGQMVDIIPPEQHAI	176				
QY	173 VPSFM---GGGKKRIHPLPYADPSLPVQPRSM DPSKDIAAYGYGSVAMKEKEMENWKQRQE	229				
Dd	177 VPSIMSGGGGGGKKRIHPLPFADPNLPIVQPRSM DPSKDIAAYGYGSVAMKEKEMGWKQKQE	236				
QY	230 RMHQTGNDGGCD-DGDDADLLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLGFFFH	288				

Db 237 RLQHVRSBGGGDWDGDDADLPLMDEARQPLSRKVPISSSSRINPYRMIIIVIRLVLGFFFH 296  
QY 289 YRVMPVNDAPALWLISVCEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDKEGQPSQ 348  
Db 297 YRVMPAKDAFALWLISVCEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDKEGQPSQ 356  
QY 349 LAPIDFVSTVDPLKEPPLVTNTVLSISVDYPVDKVSQVSDGAAMLTFEALSETSE 408  
Db 357 LAPIDFVSTVDPTKEPPLVTANTVLSISVDYPVEKVSQVSDGAAMLTFEALSETSE 416  
QY 409 FAKKWPFPCKRYNIEPRAPWYFOQKIDYLDKVAANFVRERAREREYEEFKVRINALV 468  
Db 417 FAKKWPFPCKRYNIEPRAPWYFOQKIDYLDKVAASFVRERAREREYEEFKVRINALV 476  
QY 469 AKAQKVPBEGWTMODGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNEIPRLVYVSREKRP 528  
Db 477 AKAQKVPBEGWTMODGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNEIPRLVYVSREKRP 536  
QY 529 GYNHHKKAGAMNALVRVSALVTNAPYLNDGDHYINNSKAIKEAMCFMMDPLGKKVCY 588  
Db 537 GYNHHKKAGAMNALVRVSALVTNAPYLNDGDHYINNSKAIKEAMCFMMDPLGKKVCY 596  
QY 589 VQFPQRFDGIDRHRDVRANRVVFFDINMKLDGIGPIYVGTGCVFRQALGYDAPKTK 648  
Db 597 VQFPQRFDGIDKNDVRANRVVFFDINMKLDGIGPIYVGTGCVFRQALGYDAPKTK 656  
QY 649 KPSPRTCNCPKWCFCFCCCFGNRKOKKTKTEKKLLFFKKEENSPAYALGEIDEAA 708  
Db 657 KPSPRTCNCPKWCFCFCCCFGNRKOKKTKTEKKLLFFKKEENSPAYALGEIDEAA 716  
QY 709 PGAENKAGIVNOQKLEKFFGQSSVFVTSTLLNGGTLKSASPASLKEAIIHIVISCGYED 768  
Db 717 PGADIEKAGIVNOQKLEKFFGQSSVFVTSTLLNGGTLKSASPASLKEAIIHIVISCGYED 776  
QY 769 KTDWGEKIGWIVGVTEDILTFGKMHCHGWSIYCIPIKRVAFKGSAPLNLSDRLHQVLRW 828  
Db 777 KTDWGEKIGWIVGVTEDILTFGKMHCHGWSIYCIPIKRVAFKGSAPLNLSDRLHQVLRW 836  
QY 829 ALGSIEIFFSNHCPLMYGYGGGLKFLERFSYINSIVYPWTSIPLAYCTLPAICLLTGKF 888  
Db 837 ALGSAVEIFFSKHCPLMWYGGLKFLERFSYINSIVYPWTSIPLAYCTLPAICLLTGKF 896  
QY 889 ITPELNNVASLWFMSLFCIFATSILEMWSGVGIDDWMRNEQFWIVGVSASHLFAVFG 948  
Db 897 ITPELTNVASIMFMAIFICISVTGILEMRNSGVAIDDMWRNEQFWIVGVSASHLFAVFG 956  
QY 949 LLKVIAGVDTSTVTTSKGGDDEEFSELVTRKWTLLIPTTLLLNFTGVAVGVSMAINN 1008  
Db 957 LLKVPAGIDTSTVTTSKAGDDEEFSELVTRKWTLLIPTTLLLNFTGVAVGVSMAINN 1016  
QY 1009 GYESWGPFLFGKLFPAFWIVHLYPFLKGLVGRQNRTPTIIVIVWSILASIFSLMWRIDP 1068  
Db 1017 GYESWGPFLFGKLFPAFWIVHLYPFLKGLVGRQNRTPTIIVIVWSILASIFSLMWRIDP 1076  
QY 1069 FLAKDGPFLBECGLDCN 1086  
Db 1077 FLAKSNGPFLBECGLDCN 1094  
RESULT 7  
ID AAY84118 standard; Protein; 1094 AA.  
XX AAY84118;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE Amino acid sequence of a maize cellulose synthase.  
XX  
KM Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker.  
XX

OS Zea mays.  
XX  
PN W0200009706-A2.  
XX  
PD 24-FEB-2000.  
XX  
PF 16-AUG-1999; 99WO-US18760.  
XX  
PR 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
XX  
DR WPI; 2000-224343/19.  
XX  
DR N-PSDB; AAZ99524.  
XX  
PT New genes which encode maize cellulose synthase polypeptides in plants  
PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein -  
XX  
PS Claim 15; Page 173-175; 119pp; English.  
XX  
CC The present sequence represents a maize cellulose synthase polypeptide.  
CC The cellulose synthase can be used for the improvement of stalk quality  
CC for improved stand or silage. It also provides an increased concentration  
CC of cellulose in the pericarp, hardening the kernel and improving its  
CC handling ability. The sequences are used to produce transgenic plants  
CC and seeds expressing the cellulose synthase. The polynucleotide is  
CC used for modulating, preferably increasing, the level of the synthase  
CC in a plant cell. The plants are preferably monocots. The polynucleotide  
CC is also used as a probe or primer in the detection quantitation or  
CC isolation of gene transcripts. The probes are useful in detecting  
CC deficiencies in the level of mRNA in screenings for desired transgenic  
CC plant, for detecting mutations in the gene, for monitoring upregulation  
CC of expression or changes in enzyme activity in screening assays of  
CC compounds, for detection of any number of allelic variants of the gene,  
CC or for use as molecular markers in plant breeding programs. The  
CC isolated nucleic acids of the present invention can also be used for  
CC recombinant expression of their encoded polypeptides or for use as  
CC immunogens in the preparation and/or screening of antibodies. The  
CC proteins can be employed in assays for enzyme agonists or antagonists  
CC of enzyme function or for use of immunogens or antigens to obtain  
CC antibodies specifically immunoreactive with a protein.  
XX  
SQ Sequence 1094 AA;  
Query Match 88.3%; Score 5228; DB 21; Length 1094;  
Best Local Similarity 88.7%; Pred. No. 0;  
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;  
QY 1 MEASAGLVAGSHNRNELLVIRRD--GDPGPKPPREQNGYCOICGDDVGLAPGDPFVA 57  
Db 1 MEASAGLVAGSHNRNELLVIRRDRESGAAGGAARRAEP-CQICDEVGVGFDGEFVA 59  
QY 58 CNECAPVCRDCYEYERREGTONCPQCKTRYKRLKGCQRYTGDEBEDGVDDLDFNEFN-WD 116  
Db 60 CNECAPVCRACYEYERREGSQACPCQCTRYKRLKGCPRVAGDEBEDGVDDLGEFGLOD 119  
QY 117 G----HDSQSVAESMLYGHMSYGRGDDPNGAPQAFQLNPVPLLTNGQMVDDIPPEQHAL 172  
Db 120 GAHEDDPQYVAESMLRAQMSYRGCD--AHPGFSVPVNPVPLLTNGQMVDDIPPEQHAL 176  
QY 173 VPSFM---GGGKRIHPLPYADBSLPVQPRSMPSKDLAAYGVGSVAWKERMENWKORQ 229  
Db 177 VPSYMSGGGGGKRIHPLPFADPNLPVQPRSMPSKDLAAYGVGSVAWKERMENWKORQ 236  
QY 230 RMHQTGNDGGD-DGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFH 288  
Db 237 RLQHVRSBGGGDWDGDDADLPLMDEARQPLSRKVPISSSSRINPYRMIIIVIRLVLGFFFH 296  
QY 289 YRVMPVNDAPALWLISVCEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDKEGQPSQ 348



Db 297 YRVNHPAKDAFALMLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDEKEGQPSQ 356

QY 349 LAPIDFFVSTVDPLKEPPLVTNTVLSLSVDYPVDKVSCTVSDGAAMLTFEALSETSE 408

Db 357 LAPIDFFVSTVDPTKEPPLVTANTVLSLSVDYPVEKVSCTVSDGAAMLTFEALSETSE 416

QY 409 FAKKWVPFCKRYNIEPRAPPEWYFOQKIDYLDKVAANFVRERRAMKREYEFPKRINALV 468

Db 417 FAKKWVPFSKKENIEPRAPPEWYFOQKIDYLDKVAASFVRERRAMKREYEFPKRINALV 476

QY 469 AKAQKVPEEGWTMDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVTVSREKRP 528

Db 477 AKAQKVPEEGWTMDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVTVSREKRP 536

QY 529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLLKKVCY 588

Db 537 GYNHHKKAGAMNALVRVSAVLSNAAYLLNLDCHYINNSKAIKEAMCFMMDPLLKKVCY 596

QY 589 VQFPQRFDGIDRHDRYANRNVFPDINMKGLDGIQGPITYGTGCVFRROALYGYDAPKTK 648

Db 597 VQFPQRFDGIDKNDRYANRNVFPDINMKGLDGIQGPITYGTGCVFRROALYGYDAPKTK 656

QY 649 KPPSRTCNCWPWKCFCCCCFGNRKOKTKTKTEBKULLFFPKKEENSPAYALGEIDEAA 708

Db 657 KPPSRTCNCWPWKCLSCCCSRNKNKKTKTKTEBKRLFFKKAENSPAYALGEIDEGA 716

QY 709 PGAENKAGIVNOQKLEKKFGQSSVFVSTLLENGGTLKSASPASILKEAIVHISCGYED 768

Db 717 PGADIEKAGIVNOQKLEKKFGQSSVFVASTLLENGGTLKSASPASILKEAIVHISCGYED 776

QY 769 KTDWKEIGWIYGSVTEDILTGFKMHCHGWSIYCI PKRVAFKGSAPLNLSDRLHQLRW 828

Db 777 KTDWKEIGWIYGSITEDILTGFKMHCHGWSIYCI PKRPAFKGSAPLNLSDRLHQLRW 836

QY 829 ALGSIEIFFSNHCPLMYGYGGGLKFLERFSYINSIVYPWTSIPLAYCTLPAICLLTGKF 888

Db 837 ALGSVEIFFSKICPLMYGYGGGLKFLERFSYINSIVYPWTSIPLAYCTLPAICLLTGKF 896

QY 889 ITPELNNVASLWMSLFCIFATSILEMWNGVGIDDMWRNEQFWIGVSSHLPVFGQ 948

Db 897 ITPELTNVASIWMALFICISVTGILEMRWSGVAIDDMWRNEQFWIGVSAHLFAVFGQ 956

QY 949 LLKVIAGVDTSFTVTSKGGDEEFSELYTFKWTLLIPPTLLLLNFGVAVGSNAINN 1008

Db 957 LLKVFAGIDTSFTVTSKAGDDEEFSELYTFKWTLLIPPTLLLLNFGVAVGSNAINN 1016

QY 1009 GYESWGPIFGKLFPAFWIVHLYPFLKGLVGRNRTPTIVIVWSILASIFSLMWRIDP 1068

Db 1017 GYESWGPIFGKLFPAFWIVHLYPFLKGLVGRNRTPTIVIVWSILASIFSLMWRIDP 1076

QY 1069 FLAKDGPILLECGLDN 1086

Db 1077 FLAKSNGPILLECGLDN 1094

RESULT 8

AAV58834

ID AAV58834 standard; Protein; 1165 AA.

XX AAV58834;

AC AAV58834;

XX 08-MAY-2000 (first entry)

DT 08-MAY-2000 (first entry)

DE Corn cellulose synthase (partial sequence).

XX Corn; maize; cellulose synthase; transgenic plant.

KM Corn; maize; cellulose synthase; transgenic plant.

XX Zea mays.

OS Zea mays.

XX WO200004166-A2.

PN WO200004166-A2.

XX 27-JAN-2000.

PD 27-JAN-2000.

XX

PF 13-JUL-1999; 99WO-US15871.

XX

PR 14-JUL-1998; 98US-0092844.

XX

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

XX

PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JB, Miao G;

PI Rafalski JA, Thorpe CJ;

XX

DR WPI; 2000-182431/16.

DR N-PSDB; AAZ58265.

XX

PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as

PT probes for isolating cDNAs and genes encoding homologous proteins, for

PT producing transgenic plants -

PS Claim 12; Page 45-49; 93pp; English.

XX

CC The present sequence is that of a portion of corn cellulose

CC synthase (CS), as deduced from an isolated cDNA clone (see AAV58265).

CC The invention relates to isolated CS nucleic acid fragments and to

CC CS polypeptides. It also relates to the construction of a chimeric

CC gene encoding all or a portion of the CS, in sense or antisense

CC orientation, where expression of the gene results in altered levels

CC of the CS in transformed host cells. The host cells can be used to

CC screen compounds for their ability to inhibit CS activity. CS

CC nucleic acids are also useful for producing transgenic plants having

CC altered levels of CS, and hence altered levels of fibre. CS may

CC also serve as a target for the development of novel herbicides.

XX

SQ Sequence 1165 AA;

Query Match 88.3%; Score 5228; DB 21; Length 1165;

Best local Similarity 88.7%; Pred. No. 0;

Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;

QY 1 MEASAGLVAGSHNRNELVTRD--GDPGPKPREQNGQYQICGDDVGLAFGDPFVA 57

Db 72 MEASAGLVAGSHNRNELVTRDRESGAAGGAARRAEP-CQICDEVGVGFDEPFVA 130

QY 58 CNECAFPVCRDCYEYERREGTONCPQCKTRYKRLKGCQRYTGDEBEDVDLDNEFN-WD 116

Db 131 CNECAFPVCRACYEYERREGDACPQCRTRYKRLKGCPRVAGDEBEDVDLGEFGLOD 190

QY 117 G---HDSQVAESMLYGHMSYGRGDPENGAPQAFQLNPNVPLTNGQVDDIPREQHAL 172

Db 191 GAHEDDPQYVAESMLRAQMSYGRGDP--AHFGFSPVNPVPLTNGQVDDIPREQHAL 247

QY 173 VPSFM---GGGKRIHPLPYADPSLPVQPRSMDFSKDLAAYGYSVAMKERMENWKOROE 229

Db 248 VPSYMSGGGGKRIHPLPFADBNLPVQPRSMDFSKDLAAYGYSVAMKERMENWKOROE 307

QY 230 RMHQGTGNDGGD-DGDADADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFFH 288

Db 308 RLOHVRSEGGGDWDGDADADPLMDEARQPLSRKVPISSSRINPYRMIIIRLVVLGFFFH 367

QY 289 YRVNHPVNDAFALMLISVICEIWFAMSWILDQFPKWFPIERETYLDRLSLRFDEKEGQPSQ 348

Db 368 YRVNHPAKDAFALMLISVICEIWFAMSWILDQFPKWLPIERETYLDRLSLRFDEKEGQPSQ 427

QY 349 LAPIDFFVSTVDPLKEPPLVTNTVLSLSVDYPVDKVSCTVSDGAAMLTFEALSETSE 408

Db 428 LAPIDFFVSTVDPTKEPPLVTANTVLSLSVDYPVEKVSCTVSDGAAMLTFEALSETSE 487

QY 409 FAKKWVPFCKRYNIEPRAPPEWYFOQKIDYLDKVAANFVRERRAMKREYEFPKRINALV 468

Db 488 FAKKWVPFSKKENIEPRAPPEWYFOQKIDYLDKVAASFVRERRAMKREYEFPKRINALV 547

QY 469 AKAQKVPEEGWTMDGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVTVSREKRP 528

Db 548 AKAQKVPEEGWTMDGSPWPGNNVRDHPGMIQVFLGQSGGRDVEGNELPRLVTVSREKRP 607

QY 529 GYNHHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLLKKVCY 588

Db 608 GYNHKKAGAMNALVRVSAVLNNAAYLLNLDCHYINNSKAIKEAMCFMMDPLVGKVCY 667  
QY 589 VQFPQRFIDGIDRHRANRNVVFFDINMKGLDGIGPIYVGTGCVFRQALYGVDAPKTK 648  
Db 668 VQFPQRFIDGIDKNDRYANRNVVFFDINMKGLDGIGPIYVGTGCVFRQALYGVDAPKTK 727  
QY 649 KPPSRTCNCWPKWCFCCCCGGRNKKOKTTKPKTEKKLLFFKKEENOSPAYALGEIDEAA 708  
Db 728 KPPSRTCNCWPKWCLSCCGRNKKOKTTKPKTEKKLLFFKKAENPSPAYALGEIDEGA 787  
QY 709 PGAENKAGIVNQOKLEKFGQSSVFVTSTLLENGTUKSASPASLKEAIVHISGVED 768  
Db 788 PGADIEKAGIVNQOKLEKFGQSSVFVASTLLENGTUKSASPASLKEAIVHISGVED 847  
QY 769 KTDWKEIGWIYGSVTEDILTGFKMHCHGWSIYCIIPKRVAFKGSAPLNLSDRLHQLRW 828  
Db 848 KTDWKEIGWIYGSITEDILTGFKMHCHGWSIYCIIPKRVAFKGSAPLNLSDRLHQLRW 907  
QY 829 ALGSIEIFFSNHCPLMWYGGGLKFLERFSYINSIVPWTSLPLAYCTLPATCLLTGKF 888  
Db 908 ALGSVEIFFSKHCPLMWYGGGLKFLERFSYINSIVPWTSLPLAYCTLPATCLLTGKF 967  
QY 889 ITPELNNVASLWFMSLFICIFATSILEMRSVGVIDMWRNEQFWVIGVSSHLPFAVFG 948  
Db 968 ITPELTNVASIMFMAIFICISVTGILEMRWSGVAIDMWRNEQFWVIGVSAHLFAVFG 1027  
QY 949 LLKVIAGVDTSTFTVTSKGGDDEEFSSELYTFKWTLLIPPTLLLNFIGVAVGSAINN 1008  
Db 1028 LLKVFAGIDTSFTVTSKAGDDEEFSSELYTFKWTLLIPPTLLLNFIGVAVGSAINN 1087  
QY 1009 GYESWGPPLFGKLFPAFWIVHLVLPFLKGLVGRQNRPTIIVMSILLASIFSLMWRI 1068  
Db 1088 GYESWGPPLFGKLFPAFWIVHLVLPFLKGLVGRQNRPTIIVMSILLASIFSLMWRI 1147  
QY 1069 FLAKDGPPLLEECGLDCN 1086  
Db 1148 FLAKSNGPLLEECGLDCN 1165

RESULT 9  
ABB93304

ID ABB93304 standard; Protein; 1084 AA.

XX AC ABB93304;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 2515.

KW Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms -  
XX

PS Claim 5; SEQ ID NO 2515; 261bp + Sequence listing; English.  
XX The invention relates to identifying target proteins  
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.

XX Sequence 1084 AA;

Query Match 72.3%; Score 4278.5; DB 23; Length 1084;  
Best Local Similarity 71.3%; Pred. No. 0;  
Matches 782; Conservative 127; Mismatches 161; Indels 27; Gaps 12;

QY 1 MEASAGLVAGSHRNELVIRRDGDPGPKPREQNGVCQICGDDVGLADPGDPFVACNE 60  
Db 1 MNTGRLIAGSHRNELVIRRDGDPGPKPREQNGVCQICGDDVGLADPGDPFVACNE 60  
QY 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTVG-DEEDGVDDLDNEFNMGDHD 119  
Db 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGSFRVDGDBEDIDBLEYEPD-HGMD 119  
QY 120 SQSVAESMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQMVDDIPPEQHALVPSFMG 179  
Db 120 PEHAABAALSSRLNTRGRGLDSAP-----GSQIPLLTYCDEDADMDYSDBHALIVPSTG 174  
QY 180 GGRIRHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMENWKOR-ERM---HOT 234  
Db 175 YGNRVYPAPFTDSSAPPQARSMPQKDIAYGYGSVAMKRMENWKORGEKLVYIKHEG 234  
QY 235 GNDGG---DDGDADPLMDEARQQLSKRPLPSSQINPYRMIIIRLVLGFFHYR 290  
Db 235 GNNGRGSNDDELDDPMDEGRQPLSRKLPISRINPYRMLILCRALILGLFFHYR 294  
QY 291 VMHPVNDAAFALWLSVICEIWFAMSWILDQPKWPIERETYLRLSLRDKGQPSQLA 350  
Db 295 ILHPVNDAYGLWLTSVICEIWFAMSWILDQPKWPIERETYLRLSLRDKGQPSQLA 354  
QY 351 PIDFVSTVDPLKEBPVLTNTVLSILSVDPYDVKSCYVSDDGAAMLTFEALSETSEFA 410  
Db 355 PVDVFTVSTVDPLKEBPVLTNTVLSILSVDPYDVKACVYSDDGAAMLTFEALSDTAEPFA 414  
QY 411 KKMVPFCRKNIEPRAPPEWYFOQKIDYLDKQVAANFVRRBRAMKREYEEFRVINALVAK 470  
Db 415 KKMVPFCRKNIEPRAPPEWYFQKIDYLDKQVAANFVRRBRAMKREYEEFRVINALVAT 474  
QY 471 AQKVPEEGWTMDGTMPGNNVRDHPGMIVFLGQSGGLDCBGNELPRLVYVSREKRPY 530  
Db 475 AQKVPEEGWTMDGTMPGNNVRDHPGMIVFLGHSVGRDGTGNELPRLVYVSREKRPY 534  
QY 531 NHHKKAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKVCYVQ 590  
Db 535 DHKKAGAMNSLIRVSAVLNAPYLLNLDCHYINNSKAIRSMCFMMDPSGKVCYVQ 594  
QY 591 FPQRFIDGIDRHRANRNVVFFDINMKGLDGIGPIYVGTGCVFRQALYGVDAPKTKP 650  
Db 595 FPQRFIDGIDRHRANRNVVFFDINMKGLDGIGPIYVGTGCVFRQALYGVDAPKTKP 654  
QY 651 PSRTCNCWPKWCFCCCCGGRNKKOKTTKPKTEKKLLFFKKEENOSPAYALGEIDEA--A 708  
Db 655 PGKTCNCWPKWC-CLCC-GLRKSST---KAKDKT---NTKETSQIHALENVDEGVIV 706  
QY 709 PGAENKAGIVNQOKLEKFGQSSVFVTSTLLENGTUKSASPASLKEAIVHISGVED 768  
Db 707 PVSNEKRSEATQLKEKFGQSSVFVAVASAVLONGVPRNAPACLREAIQVYISGVED 766  
QY 769 KTDWKEIGWIYGSVTEDILTGFKMHCHGWSIYCIIPKRVAFKGSAPLNLSDRLHQLRW 828  
Db 767 KTDWKEIGWIYGSVTEDILTGFKMHCHGWSIYCIIPKRVAFKGSAPLNLSDRLHQLRW 826





Db 1071 K-GGPFILEICGLDC 1083

RESULT 11

AAW33818 ID AAW33818 standard; Protein; 1084 AA.

AC AAW33818;  
DT 06-JUL-1998 (first entry)

DE Arabidopsis cellulose biosynthetic protein Ath-A.

KM Cellulose; cellulose synthase; RSW1 protein; beta-1,4-glucan;  
KW transgenic plant.

XX Arabidopsis thaliana var. Columbia.

PN WO9800549-A1.

PD 08-JAN-1998.

PF 24-JUN-1997; 97WO-AU00402.

PR 27-JUN-1996; 96AU-0000699.

PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
(AUSU ) UNIV AUSTRALIAN NAT.

PI Arioli A, Betzner AS, Peng L, Williamson RE;

DR WPI; 1998-086974/08.

DR N-PSDB; AAV06566.

PT DNA encoding cellulose biosynthetic enzyme - useful for manipulation  
of cellulose and beta-1,4-glucan

PS Claim 29; Page 130-135; 207pp; English.

XX This polypeptide is encoded by cDNA clone Ath-A (see AAV06566)  
CC that was isolated from an Arabidopsis thaliana cDNA library using  
CC PCR primers (see AAT99632-34) based on cellulose synthase RSW1  
CC genomic clone 23H12 (see AAV06563) and EST clone AAT20782 (see  
CC AAV06562). It is closely related to Arabidopsis cellulose synthase RSW1  
CC protein (see AAW33816-17). Claimed nucleic acid molecules (see  
CC AAV06562-69) coding for claimed polypeptides (see AAW33816-20 and  
CC AAW46202) involved in cellulose biosynthesis can be used to  
CC manipulate the cellulose and/or beta-glucan content of transgenic  
CC plants. Expression of nucleic acids in the sense orientation  
CC increases the level of cellulose and reduces the level of  
CC non-crystalline beta-1,4-glucan and starch, providing plants with  
CC modified strength and/or fibre properties, or having  
CC increased resistance to stresses or pests. Antisense, ribozyme or  
CC co-suppression molecules can be used to reduce the cellulose  
CC content of a transgenic plant, e.g. to improve digestibility or to  
CC alter carbon partitioning such that increased carbon is available  
CC for growth, rather than deposited as cellulose.

XX Sequence 1084 AA;

Query Match 71.8%; Score 4250.5; DB 19; Length 1084;  
Best Local Similarity 70.8%; Pred. No. 0;  
Matches 778; Conservative 128; Mismatches 162; Indels 31; Gaps 12;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPKPPRREGNGVQICGDDVGLAPGDPFVACNE 60  
Db 1 MNTGRLIAGSHNRNEFLINADESARIRSVQELSGQTCQICGDEIELTVSSSELFVACNE 60

QY 61 CAFPVCRDCYEYERREGTONCPQCKTRYKRLKGCQRYTG-DEEDGVDDLNEEFWDGHD 119  
Db 61 CAFPVCRPCYEYERREGNACPCQCKTRYKRIKSPRVGDDEEEDIDDDLEYEFD-HGMD 119

QY 120 SQSVAESMLYGHMSYGRGGDPNGAPQAFQLPQNPVPLLTNGQMVDDIPPEQHALVPSFMG 179  
Db 120 PEHAEALSSRLNTGRGGDAP-----GSQIPLLTYCDEADADMSDRHALIVPSTG 174

QY 180 GKKRIHPLPYADPSLPVQPRSMDFSKDLAAYGYSVAMKERMENWQRO-ERM---HQT 234  
Db 175 YGNRVYPAPFTDSAPPOARSMPQKDIAFYGYSVAMKDRMEVWKRQGEKLYIKHEG 234

QY 235 GNDGG---DDGDDADLPLMDEARQQLSRKPLPSSQINPYRMILIRLVVLGFFHYR 290  
Db 235 GNNGRGSNDDELDPPDMPMDEGRQPLSRKPLIRSSRINPYRMILICRLAILGLFFHYR 294

QY 291 VMHPVNDAFALMLISVCEIWFAMSWILDQFPKMPIERETYLDRLSLRFDEKGPSOLA 350  
Db 295 ILHPVNDAYGLMLTSVCEIWFASWILDQFPKMPIERETYLDRLSLRYEKEGPSOLA 354

QY 351 PIDFVSTVDPLEKPPPLVTNTVLSLSDYVVDKVSCTVSDGAAMLTFEALSETSEFA 410  
Db 355 PVDVFSTVDPLEKPPPLITANTVLSLAVDYFVDKVACVSNNGAAMLTFEALSDTADFA 414

QY 411 KKWVPFCRKYNIERRAPEMYFOQKIDYLDKVAANFVRERAMKREYERKVRINALVAK 470  
Db 415 TKWVPFCCKFNTERAPEMYFQKMDYLNKRYHAPFVRERAMKRDYERKVKINALVAT 474

QY 471 AOKVPEEGWTMDGTPWPGNNVRDHPGMIOYFLGQSGGLDCEGNEPLRLVYVSREKRP 530  
Db 475 AOKVPEERWTMDGTPWPGNNVRDHPGMIOYFLGHSVGRDITDGNELRLVYVSREKRP 534

QY 531 NHHKAGAMNALVRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQ 590  
Db 535 DHHKAGAMNSLIRVSAVLSNAPYLLNDCDHYINNSKAIRESMCFMMDPQSGKKVCYVQ 594

QY 591 FPQRFIDIRHRYANRVVFFDINMKGLDGIQPIYVGTGCVFRQALYGDAPKTKP 650  
Db 595 FPQRFIDIRHRYANRVVFFDINMKGLDGIQPIYVGTGCVFRQALYGDAPKTKP 654

QY 651 PSRTCNCWPKWPCFCCCFGNRKOKT--TRPKTEKKLLFFKKEENGSPAYALGEIDEA- 707  
Db 655 PGKTCNCWPKWC-CLCC-GLRKSKTKRADKKTNTK-----ETSKQIHALENVDEGV 704

QY 708 -APGAENEKAGIVNOQKLEKKFGQSSVFYVSTLLNGGTLSASPASLKEAIVHISCGY 766  
Db 705 IVPVSNVEKRSEATQLKEKFGQSPVFVSAVLQNGGVPRNASPACLLREAIQVISCY 764

QY 767 EDKTDWGEKEIGWYGSVTEDILTFKMGHCHGWRISYICIPKRVAFKGSAPLNSDRHQVL 826  
Db 765 QDKTEWGEKEIGWYGSVTEDILTFKMGHCHGWRISYICMPKRAAFKGSAPLNSDRHQVL 824

QY 827 RWALGSIEIFFSNHCPLWYGGGLKFLERFSYINSIVYPTWSIPLAYCTLPALCLLTG 886  
Db 825 RWALGSVEIFLSRHCPWYGGGLKFLERFSYINSVYPTWSIPLAYCSLPAYCLLTG 884

QY 887 KFITPELNNVAGIWFMSLFICIFATSIENRWSGVGIDWMRNEQFWYIGVSSHFAVF 946  
Db 885 KFIVPEISNYAGILFMLMFISIAVTGILEMGVGVGIDWMRNEQFWYIGVSSHFAVF 944

QY 947 QGLLKVIAGVDTSTFTVTSKGGDBEFSELYTFKMTLLIPTTLLLNFIIVAGVSNAI 1006  
Db 945 QGLLKVLAVNTNFTVTSKAADGAFSELYTFKMTLLIPTTLLIINIGIVGVSDAI 1004

QY 1007 NNGYESWGPIFGKLFPAFWIVHLVPLKLVGRQNRPTIVIVWSILASIFSLWRI 1066  
Db 1005 SNGYDSWGPIFGKLFPAFWIVHLVPLKLVGRQDMPTIIVWSILASIFSLWRI 1064

QY 1067 DPFLAKDDGPLLEBGLDC 1085  
Db 1065 NPFVAK-GGPVLEICGLNC 1082

RESULT 12  
ABB91831 ID ABB91831 standard; Protein; 1088 AA.



PA (FARB ) BAYER AG.  
XX  
PI Tietjen K, Weidler M;  
XX  
DR WPI; 2002-269010/31.  
XX  
PT Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms -  
XX  
PS Claim 5; SEQ ID NO 2622; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins  
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.  
XX  
SQ Sequence 1069 AA;

Query Match 71.0%; Score 4204.5; DB 23; length 1069;  
Best Local Similarity 70.6%; Pred. No. 0;  
Matches 770; Conservative 126; Mismatches 166; Indels 29; Gaps 10;

QY 1 MEASAGLVAGSHNRNELVIRRDGPGRPKPREONGOVCOICGDDVGLAPGDPFVACNE 60  
Db 1 MNTGRLIAGSHNRNEFLINADESARIRSVBELSGQTCQICGDEIELSVDSGESFVACNE 60  
QY 61 CAFPYGRDCYEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEEDGVDLIDNEFNWDGHS 120  
Db 61 CAFPYCRPCYEYERREGNOSCPOCKTRYKRIKGSPRVEGDEEDGIDDLDFEFDY---S 116  
QY 121 QSVAESMLYG--HMSYGRGGDPNGAPQAFQLPNPNVPLITNGQNVDDIPPEOHALVPSFG 178  
Db 117 RSGLESETFSRRNSEFDLASAPGS-----QIPLLTYGEEDVEIISDSHALIVSPSP 168  
QY 179 GGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQROERMQOT--G 235  
Db 169 GHIRVHQHPDPDA--AHPRPVQKDLAVYGYGSVAMKRMEEWKRKQNEKYQVVKHD 226  
QY 236 NDGGGDDDDADLPLMDEARQQLSRKIPLSSQINPYRMIIIRLVVLGFFHYRVMPV 295  
Db 227 GDSLSGDDDADIPMDDEGRQPLSRKVPKSKSKINPYRMLIVRLVILGLFFHYRIHPV 286  
QY 296 NDAFALMLISVICEIWFAMSWILDQPPKFPPIERETYLDRISLRFDEKGQPSQLAPIDFF 355  
Db 287 NDAYALMLISVICEIWFASWVLDFPKWPIERETYLDRISLRYEKEGKPSSELAGVDVF 346  
QY 356 VSTVDPLKEPPLVTTNTVLSLSVDYFVDKVSQVSDGAMLTFEALSETSEFAKKWVP 415  
Db 347 VSTVDMKEPPLITANTVLSILAVDYFVDRVACYVSDGAMLTFEALSETAEFARKWVP 406  
QY 416 FCKRYNIEPRAPEMYFOOKIDYLDKQVAANFVRERRAMKREYEFKVRINALVAKAQVP 475  
Db 407 FCKKYTEPRAPEMYFCHKMDYLNKVKHPAFVRERRAMKRDVEEFKVINALLVATAQKVP 466  
QY 476 EEGWTMQDGTWPNGNNVRDHPGMIQVFLGSGGLDCEGNELPRLVYVSREKRPYNNHKK 535  
Db 467 EEGWTMQDGTWPNGNNVRDHPGMIQVFLGNGVRDVENNELPRLVYVSREKRPFGDHHKK 526  
QY 536 AGAMNALVRVASVLTNAPYLLNDCHYINNSKAIKEAMCFMMDPLLGVKVCYVQFPQRF 595  
Db 527 AGAMNSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICYVQFPQRF 586  
QY 596 DGIIDRRYANRNVVFPDINMKGLDGIQGPYIVGTGCVFRQALYGYDAPKTKPPSRTC 655  
Db 587 DGIDKSDRYSNRNVVFPDINMKGLDGIQGPYIVGTGCVFRQALYGYDAPKTKKTKRMTC 646

QY 656 NCWPKMCFCCCCFGNRKOKTTKPTKEKKLFFKKEENOSPAYALGEIDEAPGAEN-E 714  
Db 647 NCWPKMCLFCC--GLRKNRKS--KTDDKK----KNREASKQIHAIENIEEGTKGTNDAA 698  
QY 715 KAGIVNQOKLEKKFGQSVFVSTLLNENGTLKSASPASLLKBAIHVISGVEDKTDWGK 774  
Db 699 KSPDAAQLLEKKFGQSPFVASAGMENGGLARNASPASLLREAIQVISCYEDKTEWGK 758  
QY 775 EIGMIYGSVTEIDILTFKMKHCHGWSIYICIPKRVAFKGSAPLNLSDRLHOVLRNALGSIE 834  
Db 759 EIGMIYGSVTEIDILTFKMKHSHGWSVYCTPKIPAFKGSAPINLSDRLHOVLRNALGSVE 818  
QY 835 IFFSNHCPLMWYGGGLKFLERFSYINSIVYPTWSIPLAYCTLPALICLLTGKRTPELN 894  
Db 819 IFLSRHCPIMWYGGGLKFLERLSYINSVYPTWSIFLLVYCSLPALICLLTGKRTPEIS 878  
QY 895 NYASLWFMSLFICIFATSILEMRSVSGVIGIDWWRNQFWVIGVSSHLPFAVFOGLKVIA 954  
Db 879 NYASILFMALFGSIAVTGILEMQGKVGIDDWWRNQFWVIGVSAHLFALFOGLKVIA 938  
QY 955 GVDTSFTVTSKGDDESESELYTFKWTLLIPPTLLLLNFIGVAVGSNAINGVESWG 1014  
Db 939 GVEFTNFTVTSKAADGESESELYIFKWTSLIPPTLLINVINIVGIVGSDAISNGVDSWG 998  
QY 1015 PLFGKLFFAFVIVHLVPELKLGVGRQNRPTIVYVWSILLASIFSLWVRIDPELAKDD 1074  
Db 999 PLFGRLFFAFVIVHLVPELKLGLGKQDRMPTIILVWSILLASITLLWVRVNPFAK-G 1057  
QY 1075 GPLLEEGLDC 1085  
Db 1058 GPLLEICGLDC 1068

RESULT 14  
AAB37895  
ID AAB37895 standard; Protein; 1080 AA.  
XX  
AC AAB37895;  
XX  
DT 07-MAR-2001 (first entry)  
XX  
DE Arabidopsis thaliana cellulose synthase Cels.  
XX  
KW Arabidopsis thaliana; aspen; Cela; cellulose synthase; wood;  
KW mechanical stress response element; MSRE; cellulose biosynthesis.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200071670-A2.  
XX  
PD 30-NOV-2000.  
XX  
PF 18-MAY-2000; 2000WO-US13637.  
XX  
PR 21-MAY-1999; 99US-0135280.  
XX  
PA (UNMT ) UNIV MICHIGAN TECHNOLOGICAL.  
XX  
PI Chiang VL, Wu L, Chandrasekhar JP, Carraway DT;  
XX  
DR WPI; 2001-032025/04.  
XX  
DR N-PSDB; AAC83798.  
XX  
PT Novel polynucleotide encoding cellulose synthase useful for enhancing  
PT cellulose biosynthesis, modifying lignin biosynthesis, and for  
PT accelerating growth of plants and for upregulating cellulose synthase  
PT levels -  
XX  
PS Disclosure; Fig 8; 65pp; English.  
XX  
CC The present sequence is given in a specification relating to  
CC polynucleotides encoding cellulose synthase enzymes from Populus  
CC tremuloides and Arabidopsis. The cellulose synthases are designated



CC pLcElA and AraxCelA respectively. The polynucleotides may be introduced  
CC into plants where they are useful for altering the growth of a plant.  
CC Vectors comprising a cellulose synthase promoter, or a fragment of the  
CC promoter containing one or more mechanical stress response elements  
CC (MSRE), operably linked to a cellulose synthase gene are useful for  
CC causing stress-induced gene expression in a plant cell. A cDNA encoding a  
CC protein that binds to a positive MSRE may be delivered into the plant to  
CC increase expression of cellulose synthase. A cDNA in an antisense  
CC orientation may be delivered to reduce expression of cellulose synthase.  
CC The polynucleotides enable the development of transgenic tree varieties  
CC having increased cellulose content, decreased lignin content and  
CC therefore improved wood fibre characteristics. The cellulose synthase  
CC promoters are useful for synthesising greater quantities of  
CC high-crystalline cellulose in plant and preferably in trees, which  
CC permits an increase in proportion of cellulose in transgenic plants,  
CC greater strength of juvenile wood and fiber and acceleration of overall  
CC growth rate.  
CC Note: This nucleotide sequence is given in the specification as  
CC SEQ ID NO: 5. SEQ ID NO: 5 is referred to in claim 8 but it is described  
CC as a polypeptide sequence.

XX Sequence 1080 AA;

Query Match 70.7%; Score 4185.5; DB 22; Length 1080;  
Best Local Similarity 70.7%; Pred. No. 0;  
Matches 769; Conservative 126; Mismatches 172; Indels 21; Gaps 11;

QY 7 LVAGSHRNELVIRRDGPGRPPREQNGVQICGDDVGLAPGDPFVACNECAFVVC 66  
DB 4 LIAGSHRNNEFLINADENARIRSVQELSGQTQICRDEIELTVDGEPFVACNECAFVVC 63  
QY 67 RDCYEYERREGTQNCPOCKTRYKRLKGCQRTVGDEEDGVDDLNEFNW--DGHDSQSV 124  
DB 64 RPYEYERREGNQACPOCKTRFKRLKGSPRVGEDEEDIDLDNEFEYGNNGIGFDQVS 123  
QY 125 ESMLYGHMSYG-RGGDPNGAPQAFOLNPVPLTNGQWVDIPPEQHAL-VPSFMGGGK 182  
DB 124 EGMISRRNSGPPQSDLDAPPSQ---IPLTYGDEDEVEISSDRHALIVPPSLGGHGN 179  
QY 183 RIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAWKERMENWKORQ-ERNHQTGNDGGD 241  
DB 180 RVHPVSLSDPTVAHRRLLMVPQKDLAVYGYGSVAWKDMEEWKRKQNEKLVVRHGEDPD 239  
QY 242 --DGDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMHPVND 299  
DB 240 FEDGDDADFPMMDEGRQPLSMKIPKSSKINPYRMILVLRVLGLFFHYRILHPVKDAY 299  
QY 300 ALWLISVICEIWFANSMWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLAPIDFVSTV 359  
DB 300 ALWLISVICEIWFANSMWILDQFPKWFPIERETYLDRLSLRFDKEGQPSQLSPVDVSTV 359  
QY 360 DPLKEPPLVTTNTVLSISVDYPVDKVSQVSDGAAMLTFEALSETSEFAKKWVPFCR 419  
DB 360 DPLKEPPLITANTVLSILAVDYPVDKACVSDGAAMLTFEALSETAFARKWVPFCR 419  
QY 420 YNIEPRAPFWYFOOKIDYLDKVAANFVRERRAMKREYEEFKVRINALYAKQKVPBEGW 479  
DB 420 YCIEPRAPFWYFCHKMDYLDKKNKHPAFVRERRAMKRDYEEFKVINALYATAQKVPEDGW 479  
QY 480 TMOGTPWPGNNVBDHPGMIQVFLGSGGLDCGNEELPRLVYVSREKRPQYNNHKKAGAM 539  
DB 480 TMOGTPWPGNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVSREKRPQFDHKKAGAM 539  
QY 540 NALVRVSAVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGID 599  
DB 540 NSLIRVSGVLSNAPYLLNVDCDHYINNSKALREAMCFMMDPQSGKKICVQFPQRFDGID 599  
QY 600 RHDRYARNRVVFPDINMKGDLGIQGPYVGTGCVFRQALYGYDAPRTYKPPSRTCNCP 659  
DB 600 RHDRYSNRNVVFPDINMKGDLGQGPITY-TGCVFRQALYGFDAFKKKGPRKTCNCP 658  
QY 660 KWCFCCCFGNRKQKTTTPKTEKKLLFFKKEENQSPAYALGEIDEAA--PGAENEKAG 717  
DB 660 KWCFCCCFGNRKQKTTTPKTEKKLLFFKKEENQSPAYALGEIDEAA--PGAENEKAG 717

DB 659 KWCL--LCFGSRKNRKAKTVAADKKK---KUREASKQIHALENIEBGRGHKVLNVEOST 712  
QY 718 IYNOCKLEKKFGQSSVPFVSTLLENGTLKSASPASLLKEAIIHIVISCGYEDKTDWKEIG 777  
DB 713 EAMQMKLQKKYQSPVFVASRLNENGAMARNASPACLLKEAIIQVISRGYEDKTEWKEIG 772  
QY 778 WIYGSVTEIDILTFKMHCHGWRISIYCIKRVAFKGSAPLNLSDRLHQVLRWALGSIPIFF 837  
DB 773 WIYGSVTEIDILTGSKMHSRHWYCTPKLAAFKGSAPLNLSDRLHQVLRWALGSIPIFL 832  
QY 838 SNHCPLWYGGGLKFLERESYINSIVPWTSLPLAYCTLPAICLLTGKFTPPLNVA 897  
DB 833 SRHCPIWYGGGLKWLERSYINSVVPWTSLPLIVCSLPAICLLTGKFTVPEISNYA 892  
QY 898 SLWFMSLFICIFATSLIEMRMSGVIGIDWWRNEQFVYIGVSSHFAVFOGLKRYIAGVD 957  
DB 893 SLTFMALLFSIAITGILEMQGKVGIDWWRNEQFVYIGVSAHLFALFOGLKRYIAGVD 952  
QY 958 TSFTYTSKGGDDEEFSELYTFKWTLLIPTTLILLNFIQVAVGSNAINNGYSGPPLF 1017  
DB 953 TNFTVTSKAADDEFFSDLYLFKMTSLILPMTLLIINVIGIVGVSDAISNGYDSMGPLF 1012  
QY 1018 GKLEFFAFWIVHLYPFLKGLVGRQRTPTIVIVWSILASIFSLWVRIDPFLAKDGP 1077  
DB 1013 GRLEFFALWVIHLYPFLKGLGKQDRMPTIIVWSILASILFLWVRVNPVFAK-GGPI 1071  
QY 1078 LEECGLDC 1085  
DB 1072 LEICGLDC 1079

RESULT 15  
AA58837  
ID AA58837 standard; Protein; 1039 AA.  
XX  
AC AA58837;  
XX  
DT 08-MAY-2000 (first entry)  
XX  
DE Soybean cellulose synthase (full-length).  
XX  
KW Soybean; cellulose synthase; transgenic plant.  
XX  
OS Glycine max.  
XX  
FH Key location/Qualifiers  
FT Misc-difference 201  
FT /note= "encoded by AGAA (frameshift)"  
PN WO200004166-A2.  
XX  
PD 27-JAN-2000.  
XX  
PF 13-JUL-1999; 99WO-US15871.  
XX  
PR 14-JUL-1998; 98US-0092844.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;  
PI Rafalski JA, Thorpe CJ;  
XX  
DR WPI; 2000-182431/16.  
DR N-PSDB; AA258268.  
XX  
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
PT probes for isolating cDNAs and genes encoding homologous proteins, for  
PT producing transgenic plants -  
XX  
PS Claim 6; Page 56-60; 93pp; English.  
XX  
CC The present sequence is that of full-length soybean cellulose  
CC synthase (CS), as deduced from an isolated cDNA clone (see AA588268).

CC The invention relates to isolated CS nucleic acid fragments and to  
CC CS polypeptides. It also relates to the construction of a chimeric  
CC gene encoding all or a portion of the CS, in sense or antisense  
CC orientation, where expression of the gene results in altered levels  
CC of the CS in transformed host cells. The host cells can be used to  
CC screen compounds for their ability to inhibit CS activity. CS  
CC nucleic acids are also useful for producing transgenic plants having  
CC altered levels of CS, and hence altered levels of fibre. CS may  
CC also serve as a target for the development of novel herbicides.

XX  
SQ Sequence 1039 AA;

Query Match 67.4%; Score 3986; DB 21; Length 1039;  
Best Local Similarity 67.6%; Pred. No. 0;  
Matches 738; Conservative 135; Mismatches 159; Indels 60; Gaps 14;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPKPREQNGQVCQICGDDVGLAPGDPFVACNE 60  
Db 1 MEASAGLVAGSHNRNELVIRH-GHEEPKALKNLDGQVCEICGDEVGLTVGDGLFVACNE 58  
QY 61 CAFPVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGDEEDGVDDLNEFMMDGHS 120  
Db 59 CGFPVCRPCYEYERREGSHLCPOCKTRYKRLKGSPRVEGDDEEDVDIEHEFNIDEQKN 118  
QY 121 Q--SVAESMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQ--MVDDIPPEQHALVPS 175  
Db 119 KHGQVAEAMLHGKMSYGRGPEDDNSQF-----PTPVIAGGRSPVSGEPPISSNAYGQD 173  
QY 176 FMGGG-GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQROERMQT 234  
Db 174 MLSSSLHKRVHPYVSEPG--SARWDEKKD-----GWRDMDDWKLQD----- 215  
QY 235 GNDGGGDDGDADLPIMDEARQOLSRKIPLPSSQINPYRMIIRLVVLGFFHYRVMP 294  
Db 216 GNLGPEPD-EDPDAMLEARQPLSRKPIASSKINPYRMVIVARLVILAFELRYRLMP 274  
QY 295 VNDAPALMLSVICEIMFAMSWILDQFPKPIERETYLDRLSLFDKEGQPSQIAPIDF 354  
Db 275 VHDALGLMLTSICEIMFAFSWILDQFPKFPIDRETYLDRLSIRYEREGEPMMLAPVDV 334  
QY 355 FVSTVDPLKEPPLVTNTVLSILSVYPDVKVSCYVSDGAMLTPEALSETSEFAKKMV 414  
Db 335 FVSTVDPMKEPPLVTANTVLSILAMDYPVDKISCIYISDDGASMTFBSLSETAEFAKKMV 394  
QY 415 PFCRYNIIEPRAPEWYFOQKIDYLDKQVAANFVRERAMKREYEFKVRINALVAKAQKV 474  
Db 395 PFCCKFSIEPRAPEWYFSEKIDYLDKQVPTFKERAMKREYEFKVRINALVAKAQKV 454  
QY 475 PEEGWTMODGTPWPGNNVRDHPGMIQVFLQSGGLDCEGNELPRLVYVSREKRPQYNNHK 534  
Db 455 POGGWIMODGTPWPGNNTKDPGMIQVFLQSGGLDTEGNQLPRLVYVSREKRPQYNNHK 514  
QY 535 KAGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLLKKVCYVQFPQR 594  
Db 515 KAGAMNALVRVSAVLTNAPFMLNLDCHYVNSKAREAMCFMMDPQTGKKVCYVQFPQR 574  
QY 595 FDGIDRDRYANRNVVFPDINMKGLDGIQGPITYVTGCVFRQALGYDAPKTKPPSRT 654  
Db 575 FDGIDRDRYANRNTVFPDINMKGLDGIQGPVYVTGCVFRQALGYNPPKGPKRPKMV 634  
QY 655 CNCWPKMCFCCCCFGRNRKQKTKPKTEKKLLFFKEENQSPAYALGEIDEAPGAENE 714  
Db 635 S-----CDCCPCFGSRKK-----YKEKND-----ANGEA-ASLKGMDDD 667  
QY 715 KAGIVNQKLEKKFGQSSVFVTSTLLENGTLKSASPASLKEATHVISCYEDKTDWGX 774  
Db 668 KEVLMGQNMFEKKFGQSSIFVTSTLMEEGVPSPSSPALLKEATHVISCYEDKTEWGL 727  
QY 775 EIGWYGSVTEEDILTFKMHCHGRSITICTPKRVAFKGSAPLNSDRLHQVLRWALGSIE 834  
Db 728 ELGWYGSITIEDILTGFKMHCHGRSITICMPKRAAFKGTAPINLSDRLNQVLRWALGSIE 787  
QY 835 IFFSNHCPLWYGY-GGGLKFLERFSYINSIVYPTWTSIPLLAYCTLPAICLLTGKFTTPEL 893

Db 788 IFFSHHCPLWYGFPEKKLKLMLERFAYANTTVYPTFSIPLVAYCILPAVCLITDKFIMPP 847  
QY 894 NNVASLMFMSLFICIFATSILEMKRWSGVIGDDWNRNQFWVIGVSSHLFAVPQGLKVI 953  
Db 848 STFAGLYFVALFSSIIATGILIELKWSGVSIEMWRNRNQFWVIGVSAHLFAVIGGLKVL 907  
QY 954 AGVDTSTVTSTKSGDDEEFSELYTFKWTLLIPTTLLLNFIQVAVGSNAINGYESW 1013  
Db 908 AGIDNFTVTSKATDDEEFGLYTFKWTLLIPTTLLIINIVGVAVGISDAINNGYQSW 967  
QY 1014 GPLFGKLFPAFWIVHLYPFLKGLVGRONRTPTIIVWSILLASIFSLMVRIDPELAKD 1073  
Db 968 GPLFGKLFSSFWIVHLYPFLKGLMGRONRTPTIIVWSVLLASIFSLMVRIDPEVLKT 1027  
QY 1074 DGPLLBECCGLDC 1085  
Db 1028 KGPDTKLGCINC 1039

Search completed: December 15, 2003, 13:02:49  
Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:04:25 ; Search time 36 Seconds

(without alignments)  
5610.507 Million cell updates/sec

Title: US-09-720-383C-10

Perfect score: 5918

Sequence: 1 MEASAGLVAGSHNRNMLVLI.....DPFLAKDDGPLLEECGLDCN 1086

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5918	100.0	1086	10	US-09-900-237-10 Sequence 10, Appl
2	5918	100.0	1086	12	US-10-209-059-50 Sequence 50, Appl
3	5918	100.0	1086	12	US-10-160-719-18 Sequence 18, Appl
4	5918	100.0	1086	12	US-10-160-719-38 Sequence 38, Appl
5	5570.5	94.1	1148	10	US-09-900-237-4 Sequence 4, Appl
6	5408.5	91.4	1059	12	US-10-209-059-42 Sequence 42, Appl
7	5408.5	91.4	1059	12	US-10-160-719-2 Sequence 2, Appl
8	5228	88.3	1094	12	US-10-209-059-18 Sequence 18, Appl
9	5228	88.3	1094	12	US-10-160-719-26 Sequence 26, Appl
10	5228	88.3	1094	12	US-10-160-719-46 Sequence 46, Appl
11	5228	88.3	1165	10	US-09-900-237-8 Sequence 8, Appl
12	4250.5	71.8	1084	15	US-10-229-193-8 Sequence 8, Appl
13	3991	67.4	1075	12	US-10-209-059-2 Sequence 2, Appl
14	3991	67.4	1075	12	US-10-160-719-10 Sequence 10, Appl
15	3991	67.4	1075	12	US-10-160-719-34 Sequence 34, Appl

16	3991	67.4	1075	12	US-10-160-719-54	Sequence 54, Appl
17	3986	67.4	1039	10	US-09-900-237-14	Sequence 14, Appl
18	3977.5	67.2	1074	12	US-10-209-059-46	Sequence 46, Appl
19	3977.5	67.2	1074	12	US-10-160-719-14	Sequence 14, Appl
20	3977.5	67.2	1074	12	US-10-160-719-22	Sequence 22, Appl
21	3977.5	67.2	1074	12	US-10-160-719-42	Sequence 42, Appl
22	3970.5	67.1	1091	10	US-09-900-237-26	Sequence 26, Appl
23	3945.5	66.7	1081	15	US-10-229-193-6	Sequence 6, Appl
24	3941.5	66.6	1081	15	US-10-229-193-12	Sequence 12, Appl
25	3917	66.2	1076	12	US-10-209-059-14	Sequence 14, Appl
26	3917	66.2	1076	12	US-10-160-719-58	Sequence 58, Appl
27	3905.5	66.0	1077	12	US-10-209-059-10	Sequence 10, Appl
28	3905.5	66.0	1077	12	US-10-160-719-6	Sequence 6, Appl
29	3905.5	66.0	1077	12	US-10-160-719-30	Sequence 30, Appl
30	3901	65.9	1078	12	US-10-209-059-26	Sequence 26, Appl
31	3880.5	65.6	1079	12	US-10-267-459-6	Sequence 6, Appl
32	3880.5	65.6	1079	12	US-10-209-059-22	Sequence 22, Appl
33	3876.5	65.5	1065	15	US-10-229-193-10	Sequence 10, Appl
34	3875.5	65.5	1065	10	US-09-900-237-33	Sequence 33, Appl
35	3818	64.5	1080	10	US-09-900-237-30	Sequence 30, Appl
36	3786.5	64.0	1043	12	US-10-160-719-50	Sequence 50, Appl
37	3611.5	61.0	927	12	US-10-209-059-30	Sequence 30, Appl
38	3552	60.0	881	10	US-09-838-539-8	Sequence 8, Appl
39	3521	59.5	793	10	US-09-900-237-18	Sequence 18, Appl
40	3469.5	58.6	974	10	US-09-838-539-6	Sequence 6, Appl
41	3442	58.2	1007	12	US-10-209-059-28	Sequence 28, Appl
42	3318	56.1	821	12	US-10-267-459-2	Sequence 2, Appl
43	3318	56.1	821	12	US-10-209-059-6	Sequence 6, Appl
44	3261.5	55.1	740	10	US-09-900-237-24	Sequence 24, Appl
45	2956.5	50.0	701	10	US-09-900-237-32	Sequence 32, Appl

#### ALIGNMENTS

RESULT 1  
US-09-900-237-10  
; Sequence 10, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 10  
; LENGTH: 1086  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-900-237-10

Query Match	Best Local Similarity	100.0%;	Score 5918;	DB 10;	Length 1086;	Matches 1086;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEASAGLVAGSHNRNMLVIRRDGDPKPPREONGOVCOICGDDVGLAPGDPFVACNE	60							
Db	1	MEASAGLVAGSHNRNMLVIRRDGDPKPPREONGOVCOICGDDVGLAPGDPFVACNE	60							
QY	61	CAFPVCRDCYEREREGTQNCPOCKTRYKRLKGCQRYVTGDEEDGVDDLDFNEFMWDGHS	120							
Db	61	CAFPVCRDCYEREREGTQNCPOCKTRYKRLKGCQRYVTGDEEDGVDDLDFNEFMWDGHS	120							
QY	121	QSVAESMLYGHMSYGRGDDPNGAPOAFOLNPVPLLTNGQMVDDIPPEQHALVPSFMGGC	180							



```
Db 121 QSVASMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQWDDIPPEQHALVPSFMGG 180
Qy 181 GKRHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQROERMHQTGNDGG 240
Db 181 GKRHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQROERMHQTGNDGG 240
Qy 241 DDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLGFFHYRVMPVNDAPA 300
Db 241 DDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLGFFHYRVMPVNDAPA 300
Qy 301 LMLISVCEIWFAMSWILDQFPKMFPIERETYLDRSLRFDKEGQPSQAPIDFVSTVD 360
Db 301 LMLISVCEIWFAMSWILDQFPKMFPIERETYLDRSLRFDKEGQPSQAPIDFVSTVD 360
Qy 361 PLKEPPLVTTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSEFAKKWPFCKRY 420
Db 361 PLKEPPLVTTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSEFAKKWPFCKRY 420
Qy 421 NIEPRAPEMYFOQKIDYLDKQVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480
Db 421 NIEPRAPEMYFOQKIDYLDKQVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480
Qy 481 MODGTWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNNHKKAGAMN 540
Db 481 MODGTWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNNHKKAGAMN 540
Qy 541 ALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600
Db 541 ALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600
Qy 601 HDRYANRNWVFFDINMKGLDGIQGPITYGTGCVFRQALYGYDAPKTKPPSRTCNCWPK 660
Db 601 HDRYANRNWVFFDINMKGLDGIQGPITYGTGCVFRQALYGYDAPKTKPPSRTCNCWPK 660
Qy 661 WCFCCCCFNGNRKQKTTKPKTEKKLLFFKKEENQSPAYALGEIDAAPGAENKAGIYN 720
Db 661 WCFCCCCFNGNRKQKTTKPKTEKKLLFFKKEENQSPAYALGEIDAAPGAENKAGIYN 720
Qy 721 QOKLEKKFGQSSVFVSTLLENGTLKSASPASLKEAITHVISCGYEDKTWGEIGWIT 780
Db 721 QOKLEKKFGQSSVFVSTLLENGTLKSASPASLKEAITHVISCGYEDKTWGEIGWIT 780
Qy 781 GSVTEDILTFPKMCHGWSIYCIPIKRVAFKGSAPLNSDRLHQVLRWALGSIETPSNH 840
Db 781 GSVTEDILTFPKMCHGWSIYCIPIKRVAFKGSAPLNSDRLHQVLRWALGSIETPSNH 840
Qy 841 CPLWYGGGGLKFLERFSYINSIYVWTSIPLAYCTLPACILTGKFITPELNNVASLW 900
Db 841 CPLWYGGGGLKFLERFSYINSIYVWTSIPLAYCTLPACILTGKFITPELNNVASLW 900
Qy 901 FMSLFCIFATSILEMWGSVGIIDMWRNEQFWVIGVSSHLEFAVQGLKVIAGVDTSF 960
Db 901 FMSLFCIFATSILEMWGSVGIIDMWRNEQFWVIGVSSHLEFAVQGLKVIAGVDTSF 960
Qy 961 TVTSKGGDDEEFSELYTFKWTLLIPPTLLLLNFIGVAVGSNAINNGYESWGLFGKL 1020
Db 961 TVTSKGGDDEEFSELYTFKWTLLIPPTLLLLNFIGVAVGSNAINNGYESWGLFGKL 1020
Qy 1021 FPAFWIVHLYPFLKGLVGRQNRPTIIVIVWSILLASISLWLRIDPFLAKDDGILLIE 1080
Db 1021 FPAFWIVHLYPFLKGLVGRQNRPTIIVIVWSILLASISLWLRIDPFLAKDDGILLIE 1080
Qy 1081 CGLDCN 1086
Db 1081 CGLDCN 1086
```

## RESULT 2

```
US-10-209-059-50
; Sequence 50, Application US/10209059
; Publication No. US20030163838A1
; GENERAL INFORMATION:
; APPLICANT: Dhugsa, Kanwarpal S.
```

```
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209, 059
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096, 822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371, 383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550, 483
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 1086
; TYPE: PRF
; ORGANISM: Zea mays
US-10-209-059-50
```

```
Query Match 100.0%; Score 5918; DB 12; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MEASAGLVAGSHNRNELVIRRDGDGEPKPPREQNGQVCQICGDDVGLAPGDPFVACNE 60
Db 1 MEASAGLVAGSHNRNELVIRRDGDGEPKPPREQNGQVCQICGDDVGLAPGDPFVACNE 60
Qy 61 CAFPVCRDCYERREGTQNCPOCKTRYKRLKGCQRTVGDEEDGVDDLNEFNMDGHS 120
Db 61 CAFPVCRDCYERREGTQNCPOCKTRYKRLKGCQRTVGDEEDGVDDLNEFNMDGHS 120
Qy 121 QSVASMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQWDDIPPEQHALVPSFMGG 180
Db 121 QSVASMLYGHMSYGRGDPNGAPQAFQLNPNVPLLTNGQWDDIPPEQHALVPSFMGG 180
Qy 181 GKRHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQROERMHQTGNDGG 240
Db 181 GKRHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQROERMHQTGNDGG 240
Qy 241 DDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLGFFHYRVMPVNDAPA 300
Db 241 DDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLGFFHYRVMPVNDAPA 300
Qy 301 LMLISVCEIWFAMSWILDQFPKMFPIERETYLDRSLRFDKEGQPSQAPIDFVSTVD 360
Db 301 LMLISVCEIWFAMSWILDQFPKMFPIERETYLDRSLRFDKEGQPSQAPIDFVSTVD 360
Qy 361 PLKEPPLVTTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSEFAKKWPFCKRY 420
Db 361 PLKEPPLVTTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSEFAKKWPFCKRY 420
Qy 421 NIEPRAPEMYFOQKIDYLDKQVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480
Db 421 NIEPRAPEMYFOQKIDYLDKQVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480
Qy 481 MODGTWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNNHKKAGAMN 540
Db 481 MODGTWPGNNVRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPGYNNHKKAGAMN 540
Qy 541 ALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600
Db 541 ALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGIDR 600
Qy 601 HDRYANRNWVFFDINMKGLDGIQGPITYGTGCVFRQALYGYDAPKTKPPSRTCNCWPK 660
Db 601 HDRYANRNWVFFDINMKGLDGIQGPITYGTGCVFRQALYGYDAPKTKPPSRTCNCWPK 660
Qy 661 WCFCCCCFNGNRKQKTTKPKTEKKLLFFKKEENQSPAYALGEIDAAPGAENKAGIYN 720
Db 661 WCFCCCCFNGNRKQKTTKPKTEKKLLFFKKEENQSPAYALGEIDAAPGAENKAGIYN 720
Qy 721 QOKLEKKFGQSSVFVSTLLENGTLKSASPASLKEAITHVISCGYEDKTWGEIGWIT 780
```

```
Db 721 QOKLEKFGQSSVFTSTLLNGGTLKSASPASLLKEAIIHVISCGYEDKTDWGKEIGWIY 780
Qy 781 GSVTEIDILTGFKMHCHGWSRISYICIPKRVAFKGSAPLNLSDRLHOVLRWALGSIIEIFPSNH 840
Db 781 GSVTEIDILTGFKMHCHGWSRISYICIPKRVAFKGSAPLNLSDRLHOVLRWALGSIIEIFPSNH 840
Qy 841 CPLWYGGGGLKFLERFSYINSIVYWTISIPLAYCTLPALICLLTGKFTPELNNVASLW 900
Db 841 CPLWYGGGGLKFLERFSYINSIVYWTISIPLAYCTLPALICLLTGKFTPELNNVASLW 900
Qy 901 FMSLFICIFATSIEMRWSGVIGIDWWRNEQFWVIGVSSHLPFAVFOGLLKVIAGVDTSF 960
Db 901 FMSLFICIFATSIEMRWSGVIGIDWWRNEQFWVIGVSSHLPFAVFOGLLKVIAGVDTSF 960
Qy 961 TVTSKGGDDEEFSELTYTFKWTLLIPPTLLLNFIQVAVGSNAIINNGYESWGPLFGKL 1020
Db 961 TVTSKGGDDEEFSELTYTFKWTLLIPPTLLLNFIQVAVGSNAIINNGYESWGPLFGKL 1020
Qy 1021 FFAFWIVHLYPFLKGLVGRQNRPTIIVIVWSILLASIFSLLMVRIDPFLAKDDGPLLEE 1080
Db 1021 FFAFWIVHLYPFLKGLVGRQNRPTIIVIVWSILLASIFSLLMVRIDPFLAKDDGPLLEE 1080
Qy 1081 CGLDCN 1086
Db 1081 CGLDCN 1086
```

## RESULT 3

```
US-10-160-719-18
; Sequence 18, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-18
```

```
Query Match 100.0%; Score 5918; DB 12; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEASAGLVAGSHNRNELVIRRDGPGPKPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
Db 1 MEASAGLVAGSHNRNELVIRRDGPGPKPREQNGQVCQICGDDVGLAPGGDPFVACNE 60
Qy 61 CAFPVCRDCYEREREGTONCPQCKTRYKRLKGCQRTVGDEEDGVDDLNEFNWDGHS 120
Db 61 CAFPVCRDCYEREREGTONCPQCKTRYKRLKGCQRTVGDEEDGVDDLNEFNWDGHS 120
Qy 121 QSVAESMLYGHMSYGRGDPNGAPQAFQINPNVPLLTNGQWVDIIPPEQHALVPSFMGGG 180
Db 121 QSVAESMLYGHMSYGRGDPNGAPQAFQINPNVPLLTNGQWVDIIPPEQHALVPSFMGGG 180
Qy 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQRMHQTGNDGGG 240
Db 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGYGSVAMKERMENWKQRMHQTGNDGGG 240
```

```
Qy 241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVVHPVNDABA 300
Db 241 DDGDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVVHPVNDABA 300
Qy 301 LMLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDPKEGQPSQLAPIDFVSTVD 360
Db 301 LMLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDPKEGQPSQLAPIDFVSTVD 360
Qy 361 PLKEPPLVTNTVLSISVDYPVDKVSQVSDGAMLTFEALSETSEFAKWPFCCKRY 420
Db 361 PLKEPPLVTNTVLSISVDYPVDKVSQVSDGAMLTFEALSETSEFAKWPFCCKRY 420
Qy 421 NIEPRABEWYFOQKIDVLKQVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480
Db 421 NIEPRABEWYFOQKIDVLKQVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGWT 480
Qy 481 MODGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNEIPRLVYVSREKRPQYNNHKGAGAMN 540
Db 481 MODGTPWPGNNVRDHPGMIQVFLGQSGGLDCEGNEIPRLVYVSREKRPQYNNHKGAGAMN 540
Qy 541 ALVRVSAVLTNAPYLLNDGDHYIINNSKAIKEAMCFMMDPLGKVCYVQPPQRFIDIR 600
Db 541 ALVRVSAVLTNAPYLLNDGDHYIINNSKAIKEAMCFMMDPLGKVCYVQPPQRFIDIR 600
Qy 601 HDRYANRNVPFFDINMKGLDGIQGPITYGTGCFRRQALYGYDAPKTKKPSRTCNQWPK 660
Db 601 HDRYANRNVPFFDINMKGLDGIQGPITYGTGCFRRQALYGYDAPKTKKPSRTCNQWPK 660
Qy 661 WCFCCCCFGNRKQKTKTEKTEKKLLFFKKEENQSPAYALGEIDEAPGAENEKAGIVN 720
Db 661 WCFCCCCFGNRKQKTKTEKTEKKLLFFKKEENQSPAYALGEIDEAPGAENEKAGIVN 720
Qy 721 QOKLEKFGQSSVFTSTLLNGGTLKSASPASLLKEAIIHVISCGYEDKTDWGKEIGWIY 780
Db 721 QOKLEKFGQSSVFTSTLLNGGTLKSASPASLLKEAIIHVISCGYEDKTDWGKEIGWIY 780
Qy 781 GSVTEIDILTGFKMHCHGWSRISYICIPKRVAFKGSAPLNLSDRLHOVLRWALGSIIEIFPSNH 840
Db 781 GSVTEIDILTGFKMHCHGWSRISYICIPKRVAFKGSAPLNLSDRLHOVLRWALGSIIEIFPSNH 840
Qy 841 CPLWYGGGGLKFLERFSYINSIVYWTISIPLAYCTLPALICLLTGKFTPELNNVASLW 900
Db 841 CPLWYGGGGLKFLERFSYINSIVYWTISIPLAYCTLPALICLLTGKFTPELNNVASLW 900
Qy 901 FMSLFICIFATSIEMRWSGVIGIDWWRNEQFWVIGVSSHLPFAVFOGLLKVIAGVDTSF 960
Db 901 FMSLFICIFATSIEMRWSGVIGIDWWRNEQFWVIGVSSHLPFAVFOGLLKVIAGVDTSF 960
Qy 961 TVTSKGGDDEEFSELTYTFKWTLLIPPTLLLNFIQVAVGSNAIINNGYESWGPLFGKL 1020
Db 961 TVTSKGGDDEEFSELTYTFKWTLLIPPTLLLNFIQVAVGSNAIINNGYESWGPLFGKL 1020
Qy 1021 FFAFWIVHLYPFLKGLVGRQNRPTIIVIVWSILLASIFSLLMVRIDPFLAKDDGPLLEE 1080
Db 1021 FFAFWIVHLYPFLKGLVGRQNRPTIIVIVWSILLASIFSLLMVRIDPFLAKDDGPLLEE 1080
Qy 1081 CGLDCN 1086
Db 1081 CGLDCN 1086
```

## RESULT 4

```
US-10-160-719-38
; Sequence 38, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
```

```
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-38

Query Match      100.0%; Score 5918; DB 12; Length 1086;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGQVCQICGDDVGLAPGDPFVACNE 60  
DB 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGQVCQICGDDVGLAPGDPFVACNE 60

QY 61 CAFVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGTDEEEDGVDDLNEFNMDGHS 120  
DB 61 CAFVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGTDEEEDGVDDLNEFNMDGHS 120

QY 121 QSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVDDIPPEQHALVPSFMGG 180  
DB 121 QSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVDDIPPEQHALVPSFMGG 180

QY 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAMKERMENWKQQRERHQTGNDGG 240  
DB 181 GKRIHPLPYADPSLPVQPRSMDSKDLAAYGGSVAMKERMENWKQQRERHQTGNDGG 240

QY 241 DDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMPVNDFA 300  
DB 241 DDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMPVNDFA 300

QY 301 LWLISVCEIWFAMSWILDOFPKWPPIERETYLRLSLRFDKEGQPSQAPIDFVSTVD 360  
DB 301 LWLISVCEIWFAMSWILDOFPKWPPIERETYLRLSLRFDKEGQPSQAPIDFVSTVD 360

QY 361 PLKEPPLVTNTVLSILSVDPVDKVSQVSDGAMLTFEALSETSEFAKMWPFCKRY 420  
DB 361 PLKEPPLVTNTVLSILSVDPVDKVSQVSDGAMLTFEALSETSEFAKMWPFCKRY 420

QY 421 NIEPRAPEMYFOQKIDYLDKVAANFVRERERAMKREYEEFKVRINALVAKAQKVPBEGWT 480  
DB 421 NIEPRAPEMYFOQKIDYLDKVAANFVRERERAMKREYEEFKVRINALVAKAQKVPBEGWT 480

QY 481 MODGTPWPGNNVDRHPGMIQVFLGQSGGLDCENELPRLVYVSREKRPQYNHHKAGAMN 540  
DB 481 MODGTPWPGNNVDRHPGMIQVFLGQSGGLDCENELPRLVYVSREKRPQYNHHKAGAMN 540

QY 541 ALVRSASVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDIDR 600  
DB 541 ALVRSASVLTNAPYLLNLDCDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDIDR 600

QY 601 HDRYANRNVPFDINMKGLDGIQGPITYGTGCVFRROALYGYDAPKTKKPSRTCNCPK 660  
DB 601 HDRYANRNVPFDINMKGLDGIQGPITYGTGCVFRROALYGYDAPKTKKPSRTCNCPK 660

QY 661 WCFCCCCFGNRKQKTKTKPTEKKLLFFKKEENOSPAYALGEIDEAIPAENKAGIVN 720  
DB 661 WCFCCCCFGNRKQKTKTKPTEKKLLFFKKEENOSPAYALGEIDEAIPAENKAGIVN 720

QY 721 QOKLEKFGQSSVFYVTLLENBGTLSKASAPASLKEAIIHVISCGEYEDKTWKEIGWY 780  
DB 721 QOKLEKFGQSSVFYVTLLENBGTLSKASAPASLKEAIIHVISCGEYEDKTWKEIGWY 780

QY 781 GSVTEDILTGFKMHCHGWRSTICYIPKRVAFKGSAPLNLSDRLHQLRWALGSIIEFFSNH 840  
DB 781 GSVTEDILTGFKMHCHGWRSTICYIPKRVAFKGSAPLNLSDRLHQLRWALGSIIEFFSNH 840

RESULT 5  
US-09-900-237-4  
; Sequence 4, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: B81170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 4  
; LENGTH: 1148  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-900-237-4

Query Match 94.1%; Score 5570.5; DB 10; Length 1148;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 1027; Conservative 23; Mismatches 35; Indels 5; Gaps 3;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGQVCQICGDDVGLAPGDPFVACNE 60  
DB 60 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGQVCQICGDDVGLAPGDPFVACNE 119

QY 61 CAFVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGTDEEEDGVDDLNEFNMDGHS 120  
DB 61 CAFVCRDCYEYERREGTQNCPOCKTRYKRLKGCQRTGTDEEEDGVDDLNEFNMDGHS 179

QY 120 SSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVDDIPPEQHALVPSFMGG 179  
DB 120 SSVASMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVDDIPPEQHALVPSFMGG 239

QY 180 SGLAEASMLHAHMSYGRGADLDGVPQPFHPIPNVPLLTNGQVDDIPPEQHALVPSFMGG 239  
DB 180 SGLAEASMLHAHMSYGRGADLDGVPQPFHPIPNVPLLTNGQVDDIPPEQHALVPSFMGG 299

QY 240 GDDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMPVNDFA 299  
DB 240 GDDGDDADLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMPVNDFA 359

QY 300 ALWLISVCEIWFAMSWILDOFPKWPPIERETYLRLSLRFDKEGQPSQAPIDFVSTV 359  
DB 300 ALWLISVCEIWFAMSWILDOFPKWPPIERETYLRLSLRFDKEGQPSQAPIDFVSTV 359



Db 360 ALMLISVICEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEKGPSQLAPVDFVSTV 419  
QY 360 DPLKEPPLVTTNTVLSISVDYIPVDKVSQVSDGAAMLTFEALSETSEFAKKWPFCKR 419  
Db 420 DPLKEPPLVTTNTVLSISVDYIPVDKVSQVSDGAAMLTFEALSETSEFAKKWPFCKR 479  
QY 420 YNIEPRAPEMYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGW 479  
Db 480 YSLEPRAPEMYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALVAKAQVPEEGW 539  
QY 480 TMOGTPMPGNVRDHPGMIQVFLGSGGLDCEGNEPLRLVYVSREKRPNGYNNHKKAGAM 539  
Db 540 TMOGTPMPGNVRDHPGMIQVFLGSGGHDEVEGNEPLRLVYVSREKRPNGYNNHKKAGAM 599  
QY 540 NALVRSVAVLTNAPYLLNDGDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGID 599  
Db 600 NALVRSVAVLTNAPYLLNDGDHYINNSKAIKEAMCFMMDPLGKKVCYVQFPQRFDGID 659  
QY 600 RHDYRANRVVFFDINMKGLDGIQGPITYVGTGCVFRQALYGYDAPKTKKPPSRTCNCWP 659  
Db 660 RHDYRANRVVFFDINMKGLDGIQGPITYVGTGCVFRQALYGYDAPKTKKPPSRTCNCWP 719  
QY 660 KWCFCGCCCFGNRKOK--KTKPKTEKKLLFFKKEENOSPAYALGEIDEAAPGAENKA 716  
Db 720 KWCICCCCFGNRKTKKTKTSKPKFEKIKKL-FKKKENQAPAYALGEIDEAAPGAENKA 778  
QY 717 GIVNQQLEKKFGQSSVFTSTLLENGGTLKSASPASLLKEAIIHVISCGYEDKTDWKEI 776  
Db 779 SIVNQQLEKKFGQSSVFTSTLLENGGTLKSASPASLLKEAIIHVISCGYEDKTDWKGDI 838  
QY 777 GVIYGSVTEIDILTFKMHGHWRSIYCIPIKRAAFKGSAPLNSDRLHQVLRWALGSIEIF 836  
Db 839 GVIYGSVTEIDILTFKMHGHWRSIYCIPIKRAAFKGSAPLNSDRLHQVLRWALGSIEIF 898  
QY 837 FSNHCPLMYGGGGLKFLERFSYINSIVYPTWTSIPLAYCTLPACLLTGKFTPELNNV 896  
Db 899 FSNHCPLMYGGGGLKFLERFSYINSIVYPTWTSIPLAYCTLPACLLTGKFTPELNNV 958  
QY 897 ASLMFMSLFICIFATSIENKWSGVGIDDMWRNEQFWIIGVSSHFAVFGGLKVYAGV 956  
Db 959 ASLMFMSLFICIFATSIENKWSGVGIDDMWRNEQFWIIGVSSHFAVFGGLKVYAGV 1018  
QY 957 DTSFTVTSKGGDDEEFSELYTFKWTLLIPTTLILLNFIVGAVGSNAINNGYESWGPL 1016  
Db 1019 DTSFTVTSKGGDDEEFSELYTFKWTLLIPTTLILLNFIVGAVGSNAINNGYESWGPL 1078  
QY 1017 FGKLPFAFVIVHLYPFLKGLVGRQNRPTIIVIVWSILLASISFLMWRIDPFLAKDGP 1076  
Db 1079 FGKLPFAFVIVHLYPFLKGLVGRQNRPTIIVIVWSILLASISFLMWRIDPFLAKDGP 1138  
QY 1077 LLEEGLDCN 1086  
Db 1139 LLEEGLDCN 1148

RESULT 6  
US-10-209-059-42  
; Sequence 42, Application US/10209059  
; Publication No. US20030163838A1  
; GENERAL INFORMATION:  
; APPLICANT: Dhugga, Kanwarpal S.  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: thereof  
; FILE REFERENCE: 0864R2  
; CURRENT APPLICATION NUMBER: US/10/209, 059  
; PRIOR FILING DATE: 2002-07-31  
; PRIOR APPLICATION NUMBER: 60/096, 822  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 09/371, 383  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR APPLICATION NUMBER: 09/550, 483  
; PRIOR FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42  
; LENGTH: 1059  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-209-059-42

Query Match 91.4%; Score 5408.5; DB 12; Length 1059;  
Best Local Similarity 94.0%; Pred. No. 0;  
Matches 995; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

QY 33 EQNGQVCQICGDDVGLAPGDPFVACNEGAPVVCRCDCYERREBTONCPQCKTRYKRLK 92  
Db 3 QNRGQVCQICGDDVGRNPDGEPFVACNECAPPICRDCYERREBTONCPQCKTRFKRFK 62  
QY 93 GCQRVTGDEEDGVDDLNEFNW-DGHSQVAESMLYGHMSYGRGDPNGAPQAFQLPNP 151  
Db 63 GCARVPGEDEEDGVDDLNEFNWSDKHSQYLABSMLHAHMSYGRGADLDGVPQFHPHP 122  
QY 152 NVPLLTNGQWVDIPREQHVLVPSFMGGGKRHPPLPYADPSLPVQPRSMDBSKDLAYG 211  
Db 123 NVPLLTNGQWVDIPRDQHALVPSFVGGGKRHPPLPYADPNLPVQPRSMDBSKDLAYG 182  
QY 212 YGSVAMKEREMWKKQROERHQTGNDGGDDGDDADLPLMDEARQULSRKIPLPSSQINP 271  
Db 183 YGSVAMKEREMWKKQROERHQTGNDGGDDGDDADLPLMDEARQULSRKIPLPSSQINP 242  
QY 272 YRMIIIRLVVLGFFHFHYRVMPVNDAPALMLISVICEIWFAMSWILDQFPKWPPIERET 331  
Db 243 YRMIIIRLVVLGFFHFHYRVMPVNDAPALMLISVICEIWFAMSWILDQFPKWPPIERET 302  
QY 332 YLDRLSLRFDEKGPSQLAPIDFVFSVTDPLKEPPLVTTNTVLSISVDYIPVDKVSQV 391  
Db 303 YLDRLSLRFDEKGPSQLAPVDFVFSVTDPLKEPPLVTTNTVLSISVDYIPVDKVSQV 362  
QY 392 DDGAAMLTFEALSETSEFAKKWPFCKRYNIEPRAPEMYFOQKIDYLDKVAANFVRER 451  
Db 363 DDGAAMLTFEALSETSEFAKKWPFCKRYNIEPRAPEMYFOQKIDYLDKVAANFVRER 422  
QY 452 AMKREYEEFKVRINALVAKAQVPEEGWTMOGTPMPGNVRDHPGMIQVFLGSGGLDC 511  
Db 423 AMKREYEEFKVRINALVAKAQVPEEGWTMOGTPMPGNVRDHPGMIQVFLGSGGHV 482  
QY 512 EGNEPLRLVYVSREKRPNGYNNHKKAGANALVRSVAVLTNAPYLLNDGDHYINNSKAIK 571  
Db 483 EGNEPLRLVYVSREKRPNGYNNHKKAGANALVRSVAVLTNAPYLLNDGDHYINNSKAIK 542  
QY 572 EAMCFMMDPLGKKVCYVQFPQRFDGIDRHDYRANRVVFFDINMKGLDGIQGPITYGTG 631  
Db 543 EAMCFMMDPLGKKVCYVQFPQRFDGIDRHDYRANRVVFFDINMKGLDGIQGPITYGTG 602  
QY 632 CVFRQALYGYDAPKTKKPPSRTCNCWPKCFCCCFGNRKOK--KTKPKTEKKLLFF 688  
Db 603 CVFRQALYGYDAPKTKKPPSRTCNCWPKCICCCCFGNRKTKKTKTSKPKFEKIKKL- 661  
QY 689 FKKEENOSPAYALGEIDEAAPGAENKACIVNQQLEKKFGQSSVFTSTLLENGGTLKS 748  
Db 662 FKKEENQAPAYALGEIDEAAPGAENKASIVNQQLEKKFGQSSVFTSTLLENGGTLKS 721  
QY 749 ASPASLKEAIIHVISCGYEDKTDWKEIGVIYGSVTEIDILTFKMHGHWRSIYCIPIKRV 808  
Db 722 ASPASLKEAIIHVISCGYEDKTDWKEIGVIYGSVTEIDILTFKMHGHWRSIYCIPIKRA 781  
QY 809 AFKGSAPLNSDRLHQVLRWALGSIEIFFSNHCPLMYGGGGLKFLERFSYINSIVYPTW 868  
Db 782 AFKGSAPLNSDRFHQVLRWALGSIEIFFSNHCPLMYGGGGLKFLERFSYINSIVYPTW 841  
QY 869 SIPLAYCTLPACLLTGKFTPELNNVVASLMFMSLFICIFATSIENKWSGVGIDDMWR 928  
Db 842 SIPLAYCTLPACLLTGKFTPELNNVVASLMFMSLFICIFATSIENKWSGVGIDDMWR 901  
QY 929 NEQFWIIGVSSHFAVFGGLKVYAGVDTSTFTVTSKGGDDEEFSELYTFKWTLLIPT 988

Db 902 NEOFVWIGVSSHLFAVFOGLKVIAGVDTSTFTVTSKGGDDEESELTYTFKWTLLIPT 961  
Qy 989 TLLLNFIGVAVGSNAINNGYESWGPFGKLFPAFWIVHLYPEPLKGLVGRONRTPTIV 1048  
Db 962 TLLLNFIGVAGISNAINNGYESWGPFGKLFPAFWIVHLYPEPLKGLVGRONRTPTIV 1021  
Qy 1049 IVWSILASIFSLLWVRIDPFLAKDDGPLLECGLDN 1086  
Db 1022 IVWSILASIFSLLWVRIDPFLAKDDGPLLECGLDN 1059

RESULT 7

US-10-160-719-2  
; Sequence 2, Application US/10160719  
; Publication No. US20030167528A1  
; GENERAL INFORMATION:  
; APPLICANT: Kanwarpal S. Dhugga  
; APPLICANT: Timothy G. Helentjaris  
; APPLICANT: Benjamin A. Bowen  
; APPLICANT: Xun Wang  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 0864  
; CURRENT APPLICATION NUMBER: US/10/160,719  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/371,383A  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR APPLICATION NUMBER: US 60/096,822  
; PRIOR FILING DATE: 1998-08-17  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1059  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-160-719-2

Query Match 91.4%; Score 5408.5; DB 12; Length 1059;  
Best Local Similarity 94.0%; Pred. No. 0;  
Matches 995; Conservative 22; Mismatches 36; Indels 5; Gaps 3;

Qy 33 EQNGOVCOICGDDVGLAPGGDPFVACNECAFVCRDCYEREREGTONCPOCKTRYKRLX 92  
Db 3 QNRGQVCOICGDDVGRNPDGEPFVACNECAFPCRDCEYERREGTONCPOCKTRYKRFK 62  
Qy 93 GCQRVTGDEEEDGVDDLNEFNW-DGHSQSVASMLYGHMSYGRGGDPNGAPQAQOLNP 151  
Db 63 GCARVPGEDEEDGVDDLNEFNWSDKHSQYLAESMLHAHMSYGRGADGVQPFHPPI 122  
Qy 152 NVPLLTNGQWVDIPPEQHALVPSFMGGGKRIHPLPYADPSLPVQPRSMDSKDLAAYG 211  
Db 123 NVPLLTNGQWVDIPPEQHALVPSFVGGGKRIHPLPYADPNLPVQPRSMDSKDLAAYG 182  
Qy 212 YGSVAMKERMENWQORERHQTGNDGGDDADLPLMDEARQQLSRKIPLPSSQINP 271  
Db 183 YGSVAMKERMESWQORERHQTNDGGDDADLPLMDEARQPLSRKIPLPSSQINP 242  
Qy 272 YRMIIIRLVVLGFFHYRVMPVNDAPALWLVSVCEIWFAMSWILDQFPKMPPIERET 331  
Db 243 YRMIIIRLVLCFFHYRVMPVNDAPALWLVSVCEIWFAMSWILDQFPKMPPIERET 302  
Qy 332 YLDRLSLRFDKEGQPSQAPIDFVSTVDPLKEPPLVTNTVLSILSVDPYDKVSCYVS 391  
Db 303 YLDRLSLRFDKEGHPSQLAPVDFVSTVDPLKEPPLVTANTVLSILSVDPYDKVSCYVS 362  
Qy 392 DDGAAMLTFEALSETSEFAKKWVPFCRKYNIERAPAEWYFQOKIDYLDKDYAANFVRER 451  
Db 363 DDGAAMLTFEALSETSEFAKKWVPFCRKYSLERAPAEWYFQOKIDYLDKDYAANFVRER 422  
Qy 452 AMKREYEFPKVRINALVAKAQKVPBEGWTMDGTPWPGNNVRDHPGMIQVFLGQSGGLDC 511  
Db 423 AMKREYEFPKVRINALVAKAQKVPBEGWTMDGTPWPGNNVRDHPGMIQVFLGQSGGHDV 482

Qy 512 EGNELPRLVYVSREKRPQGVNHHKAGAMNALVRVSAVLTNAPYLLNDLDCDHYINNSKAIK 571  
Db 483 EGNELPRLVYVSREKRPQGVNHHKAGAMNALVRVSAVLTNAPYLLNDLDCDHYINNSKAIK 542  
Qy 572 EAMCFMMDPLGKVCYVQFPQRFDDGIDRHRDRIANRVVFPDINMGLDGIQPIYVGTG 631  
Db 543 EAMCFMMDPLGKVCYVQFPQRFDDGIDRHRDRIANRVVFPDINMGLDGIQPIYVGTG 602  
Qy 632 CVFRRQALYGYPKTKKPSRTCNCPKWCFCCECFGNRKOK---KTKPKTEKKLLF 688  
Db 603 CVFRRQALYGYPKTKKPSRTCNCPKWCICCCCFGNRKTKKTKTSKPKFEKIKKL- 661  
Qy 689 FKKEENQSPAYALGEIDEAPGAENEKAGIVNQOKLEKKFGQSSVFTSTLLENGTLLKS 748  
Db 662 FKKEENQAPAYALGEIDEAPGAENEKASIVNQOKLEKKFGQSSVFASTLLENGTLLKS 721  
Qy 749 ASPASLLKEAIVHISCGEYEDKTDWKEIGWTYGSVTEIDILTGFKMHCHGWSIYCIIPKRV 808  
Db 722 ASPASLLKEAIVHISCGEYEDKTDWKEIGWTYGSVTEIDILTGFKMHCHGWSIYCIIPKRA 781  
Qy 809 AFKGSAPLNLSDRLHQVLRNALGSIIEFPSNHCPLWYGGGLKFLERFSYINSIYIPWT 868  
Db 782 AFKGSAPLNLSDRFHQVLRNALGSIIEIFSNHCPLWYGGGLKFLERFSYINSIYIPWT 841  
Qy 869 SIPLIAYCTLPALCLLTGKFTPELNNVASLWMSLFICIFATSILEMRWSGVGIDMWR 928  
Db 842 SIPLIAYCTLPALCLLTGKFTPELNNVASLWMSLFICIFATSILEMRWSGVGIDMWR 901  
Qy 929 NEOFVWIGVSSHLFAVFOGLKVIAGVDTSTFTVTSKGGDDEESELTYTFKWTLLIPT 988  
Db 902 NEOFVWIGVSSHLFAVFOGLKVIAGVDTSTFTVTSKGGDDEESELTYTFKWTLLIPT 961  
Qy 989 TLLLNFIGVAVGSNAINNGYESWGPFGKLFPAFWIVHLYPEPLKGLVGRONRTPTIV 1048  
Db 962 TLLLNFIGVAGISNAINNGYESWGPFGKLFPAFWIVHLYPEPLKGLVGRONRTPTIV 1021  
Qy 1049 IVWSILASIFSLLWVRIDPFLAKDDGPLLECGLDN 1086  
Db 1022 IVWSILASIFSLLWVRIDPFLAKDDGPLLECGLDN 1059

RESULT 8

US-10-209-059-18  
; Sequence 18, Application US/10209059  
; Publication No. US20030163838A1  
; GENERAL INFORMATION:  
; APPLICANT: Dhugga, Kanwarpal S.  
; APPLICANT: Wang, Haiyin  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 0864R2  
; CURRENT APPLICATION NUMBER: US/10/209,059  
; PRIOR FILING DATE: 2002-07-31  
; PRIOR APPLICATION NUMBER: 60/096,822  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 09/371,383  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR APPLICATION NUMBER: 09/550,483  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 1094  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-209-059-18

Query Match 88.3%; Score 5228; DB 12; Length 1094;  
Best Local Similarity 88.7%; Pred. No. 0;  
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;  
Qy 1 MEASAGLVAGSHNRNELVVIIRD---GDFGPKPREQNGVCOICGDDVGLAPGGDPFVA 57

Db 1 MEASAGLVAGSHNRNELVIRDRRESGAAGGAARAEAP-CQICGDEVGVFDGEFPVA 59  
Qy 58 CNECAPVPCRDCEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEBDGVDDLNEFN-WD 116  
Db 60 CNECAPVPCRDCEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEBDGVDDLNEFN-WD 119  
Qy 117 G---HDSQSVAESMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVDDIPPEQHAL 172  
Db 120 GAAHEDDPQYVAESMLRAQMSYGRGD---AHPGFSVPVNPVPLLTNGQVDDIPPEQHAL 176  
Qy 173 VPSFM---GGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMENWKOROE 229  
Db 177 VPSYMSGGGGGKRIHPLPFADPNLPVQPRSMDSKDLAAYGYSVAMKERMENWKOROE 236  
Qy 230 RMEQTGNDGGD-DGDDADPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFH 288  
Db 237 RLOHVRSEGGGDWDGDDADPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFH 296  
Qy 289 YRVMPVNDAPALWLVSVCEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEKQPSQ 348  
Db 297 YRVMPVNDAPALWLVSVCEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEKQPSQ 356  
Qy 349 LAPIDFVSTVDPLKEPPLVTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSE 408  
Db 357 LAPIDFVSTVDPTKEPPLVTANTVLSILSVDPVEKVSQVSDGAAMLTFEALSETSE 416  
Qy 409 FAKKWPVFCRKRYNIEPRAPEMYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALV 468  
Db 417 FAKKWPVFCRKRYNIEPRAPEMYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALV 476  
Qy 469 AKAQVPEEGWTMODGTPWPGNNVRDHGMIQVFLGQSGGLDCGNEPLRLVYVSREKRP 528  
Db 477 AKAQVPEEGWTMODGTPWPGNNVRDHGMIQVFLGQSGGLDCGNEPLRLVYVSREKRP 536  
Qy 529 GYNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCY 588  
Db 537 GYNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCY 596  
Qy 589 VQFPQRFQDIDRDRYANRVVFPDINMGLDGIQGPVYVGTGCVFRQALYGYDAPKTX 648  
Db 597 VQFPQRFQDIDRDRYANRVVFPDINMGLDGIQGPVYVGTGCVFRQALYGYDAPKTX 656  
Qy 649 KPPSRTCNCPKWCFCFCCCFGNRKQCKTKRPTKEKKLLFFKKEENQSPAYALGEIDEAA 708  
Db 657 KPPSRTCNCPKWCFCFCCCFGNRKQCKTKRPTKEKKLLFFKKEENQSPAYALGEIDEAA 716  
Qy 709 PGAENKAGIVNQQLEKKFGQSSVFTSTLLENGGTLKSASPASLKEAIIHVISCGYED 768  
Db 717 PGAENKAGIVNQQLEKKFGQSSVFTSTLLENGGTLKSASPASLKEAIIHVISCGYED 776  
Qy 769 KTDWGEIGWIVGVTEDILTGFKMHCHGWSIYCI PKRPAFKGSAPLNLSDRLHQLRW 828  
Db 777 KTDWGEIGWIVGVTEDILTGFKMHCHGWSIYCI PKRPAFKGSAPLNLSDRLHQLRW 836  
Qy 829 ALGSIIEFNSHCPLMYGGGLKFLERSYINSIVYPWTSIPLAAYCTLPAICLLTGKF 888  
Db 837 ALGSIIEFNSHCPLMYGGGLKFLERSYINSIVYPWTSIPLAAYCTLPAICLLTGKF 896  
Qy 889 ITPELNNVASLWMSLFCIFATSILEMWMSGVGIDDMWRNEQFWIVGVSASHLFAVFOG 948  
Db 897 ITPELNNVASLWMSLFCIFATSILEMWMSGVGIDDMWRNEQFWIVGVSASHLFAVFOG 956  
Qy 949 LKVIAGVDTSTVTYSGGDDSEFSELYTFKTTLLIPTTLLLNFIQVAVGSNAIIN 1008  
Db 957 LKVIAGVDTSTVTYSGGDDSEFSELYTFKTTLLIPTTLLLNFIQVAVGSNAIIN 1016  
Qy 1009 GYESWGPLEFGLFFAFWVIVHLYPFLKGLVGRQNRPTPTIVIVMSILASISFLMWRIDP 1068  
Db 1017 GYESWGPLEFGLFFAFWVIVHLYPFLKGLVGRQNRPTPTIVIVMSILASISFLMWRIDP 1076  
Qy 1069 FLAKDDGPLLEECGLDCN 1086

Db 1077 FLAKSNGPLLEECGLDCN 1094  
RESULT 9  
US-10-160-719-26  
; Sequence 26, Application US/10160719  
; Publication No. US20030167528A1  
; GENERAL INFORMATION:  
; APPLICANT: Kanwarpal S. Dhugga  
; APPLICANT: Timothy G. Helentjaris  
; APPLICANT: Benjamin A. Bowen  
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses  
; TITLE OF INVENTION: thereof  
; FILE REFERENCE: 0864  
; CURRENT FILING DATE: US/10/160, 719  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR FILING DATE: US/09/371, 383A  
; PRIOR FILING DATE: 1999-08-06  
; PRIOR FILING DATE: 1998-08-17  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 1094  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-160-719-26  
Query Match 88.3%; Score 5228; DB 12; Length 1094;  
Best Local Similarity 88.7%; Pred. No. 0;  
Matches 974; Conservative 38; Mismatches 70; Indels 16; Gaps 7;  
Qy 1 MEASAGLVAGSHNRNELVIRDR--GDPGPKPPREONQOVQICGDDVGLAPGGEFPVA 57  
Db 1 MEASAGLVAGSHNRNELVIRDRRESGAAGGAARAEAP-CQICGDEVGVFDGEFPVA 59  
Qy 58 CNECAPVPCRDCEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEBDGVDDLNEFN-WD 116  
Db 60 CNECAPVPCRDCEYERREGTQNCPOCKTRYKRLKGCQRVTGDEEBDGVDDLNEFN-WD 119  
Qy 117 G---HDSQSVAESMLYGHMSYGRGDPNGAPQAFQNLNPNVPLLTNGQVDDIPPEQHAL 172  
Db 120 GAAHEDDPQYVAESMLRAQMSYGRGD---AHPGFSVPVNPVPLLTNGQVDDIPPEQHAL 176  
Qy 173 VPSFM---GGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMENWKOROE 229  
Db 177 VPSYMSGGGGGKRIHPLPFADPNLPVQPRSMDSKDLAAYGYSVAMKERMENWKOROE 236  
Qy 230 RMEQTGNDGGD-DGDDADPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFH 288  
Db 237 RLOHVRSEGGGDWDGDDADPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVVLGFFH 296  
Qy 289 YRVMPVNDAPALWLVSVCEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEKQPSQ 348  
Db 297 YRVMPVNDAPALWLVSVCEIWFAMSWILDQFPKWPPIERETYLDRLSLRFDEKQPSQ 356  
Qy 349 LAPIDFVSTVDPLKEPPLVTNTVLSILSVDPVDKVSQVSDGAAMLTFEALSETSE 408  
Db 357 LAPIDFVSTVDPTKEPPLVTANTVLSILSVDPVEKVSQVSDGAAMLTFEALSETSE 416  
Qy 409 FAKKWPVFCRKRYNIEPRAPEMYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALV 468  
Db 417 FAKKWPVFCRKRYNIEPRAPEMYFOQKIDYLDKVAANFVRERRAMKREYEEFKVRINALV 476  
Qy 469 AKAQVPEEGWTMODGTPWPGNNVRDHGMIQVFLGQSGGLDCGNEPLRLVYVSREKRP 528  
Db 477 AKAQVPEEGWTMODGTPWPGNNVRDHGMIQVFLGQSGGLDCGNEPLRLVYVSREKRP 536  
Qy 529 GYNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCY 588  
Db 537 GYNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKKVCY 596



QY	589	VQFPQRFDDIDRHDRYANRNVVFPDINMKGLDGIQGPYVGTGCVFRROALYGDAPKTK	648
Db	597	VQFPQRFDDIDKNDRYANRNVVFPDINMKGLDGIQGPYVGTGCVFRROALYGDAPKTK	656
QY	649	KPSPRTCNCPKWCFCFCCCFGNRKQKTKTKPTEKKKLLFFKKEENOSPAYALGEIDEAA	708
Db	657	KPSPRTCNCPKWCFLSCCSCRNNKKKTKTKPTEKKKLLFFKKAENSPAYALGEIDEGA	716
QY	709	PGAENKAGIVNQQKLEKKEGQSSVFVTSTLLENGITLKASPASLLKEAIIHVISGVED	768
Db	717	PGADIEKAGIVNQQKLEKKEFGQSSVFVASTLLENGITLKASPASLLKEAIIHVISGVED	776
QY	769	KTDWGKEIGMIYGSVTEBILITGFKMHCHGWRISIYCIKRYAFKGSAPLNLSDRLHOVLRW	828
Db	777	KTDWGKEIGMIYGSITEBILITGFKMHCHGWRISIYCIKRPAPFKGSAPLNLSDRLHOVLRW	836
QY	829	ALGSEIEFFSNHCPLWYGGGLKFLERFSYINSIYVPWTSIPLLAYCTLPALCILLTGKF	888
Db	837	ALGSVEIFFSKHCPMLWYGGGLKFLERFSYINSIYVPWTSIPLLAYCTLPALCILLTGKF	896
QY	889	ITPELNVASLWFMSLFCIFATSILEMRRMSGVIGIDDMWRNEQFWIVIGVSSHFAVFOG	948
Db	897	ITPELNVASIMFMALFICISVTGILEMRMSGVAIDDMWRNEQFWIVIGVSAHLFAVFOG	956
QY	949	LLKVIAGVDTSFVTYTSKGGDEEFSELYTFKWTTLIPPTLLLLNFIQVAVAGVSNAINN	1008
Db	957	LLKVFAGIDTSFTVYTSKAGDEEFSELYTFKWTTLIPPTLLLLNFIQVAVAGISNAINN	1016
QY	1009	GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRONRPTPIVMSILLASIFSLMWRIDP	1068
Db	1017	GYESWGPLFGKLFFAFWVIVHLYPFLKGLVGRONRPTPIVMSILLASIFSLMWVAVDP	1076
QY	1069	FLAKDDGPLLEBCGLDCN 1086	
Db	1077	FLAKSNGPLLEBCGLDCN 1094	

```

RESULT 10
US-10-160-719-46
; Sequence 46, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-46

```

	Query Match	88.3%;	Score 5228;	DB 12;	Length 1094;	
	Best Local Similarity	88.7%;	Pred. No. 0;			
	Matches 974;	Conservative 38;	Mismatches 70;	Indels 16;	Gaps 7;	
QY	1 MEASAGLVAGSHNRNELVVRIRD--GDGPKPPREQNGVCQICDDVGLAPGGDPFVA			:	:	57
Dd	1 MEXSAGLVAGSHNRNELVVRIRDRESGAAGGAARRAEAP-CQICGDEVGVGFDEGPFA			:	:	59
QY	58 CNECAFVPCRDCEYERREGTONPQCKTRYKRLKGQRVTGDHEEDGVDDLDFNEFN-WD		:	:	:	116

Db	60	CNECAFVCRACYEYERREGSQACPQCRTYKRLKGCPRVAGDEEBDGVDDLEGEFGLQD	119
QY	117	G---HDSQVAESMLYGHMSYGRGGDPNGAPQAFQLNPNVPLLTNGQVDDIPPEQHAL	172
Db	120	GAAHEDDPQYVAESMLRAQMSYGRGGD--AHGFSPVPNVPLLTNGQVDDIPPEQHAL	176
QY	173	VPSFM--GGGKRIRHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMENWKORQ	229
Db	177	VPSYMSGGGGGKKRIHPLPFADPNLPVQPRSMDSKDLAAYGYSVAMKERMEGKQOE	236
QY	230	RMHQTNDDGGD-DGDDADPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVYLGFEFH	288
Db	237	RLQHVRSSEGGDMDGDDADPLMDEARQPLSRKVPISSSRINPYRMITIVRLVYLGFEFH	296
QY	289	YRVMPVNDAPALMLISVICEIMFAMSWILDPFKMFPIERETYLDRLSLRFDEGQPSQ	348
Db	297	YRVMPAKDAFALMLISVICEIMFAMSWILDPFKMLPIERETYLDRLSLRFDEGQPSQ	356
QY	349	LAPIDFVSTVDPLKEPPLVTNTVLISLSDVPDVKVSCYSDDGAAMLTFEALSETSE	408
Db	357	LAPIDFVSTVDPTKEPPLVTANTVLISLSDVPYEVKVCYSDDGAAMLTFEALSETSE	416
QY	409	FAKKWVPFCRKNIEPRAPBEMWFQOKIDYLDKQVAANFVRBRAMKREYEEFKVRINALV	468
Db	417	FAKKWVPFSKKFNIERAPBEMWFQOKIDYLDKQVAASFVRBRAMKREYEEFKVRINALV	476
QY	469	AKAQVPEEGWTMDGTWPGNNVRDHPGMIQVFLQSGSGGLDCEGNELPRLVYVSREKRP	528
Db	477	AKAQVPEEGWTMDGSPWPGNNVRDHPGMIQVFLQSGSGRDEGNELPRLVYVSREKRP	536
QY	529	GYNHKKAGAMNALYRVSAYLTNPAYLLNLDCHYINNSKAIKEAMCFMMDPLVGKKVCY	588
Db	537	GYNHKKAGAMNALYRVSAYLSNAYLLNLDCHYINNSKAIKEAMCFMMDPLVGKKVCY	596
QY	589	VQFPQRFDDIRHRYANRNVPFDDINMKGLDGIQGPITYGTGCVFRQALYGYDAPKTK	648
Db	597	VQFPQRFDDIKNDRYANRNVPFDDINMKGLDGIQGPITYGTGCVFRQALYGYDAPKTK	656
QY	649	KPPSRTCNCWPKWCFFCCCCFGNRKQKTKPKTEKKLLPFKKEENQSPAYALGEIDEAA	708
Db	657	KPPSRTCNCWPKWCLSCCCSRNKNKTKTKPKTEKKRLPFKKAENPSPAYALGEIDEGA	716
QY	709	PGAENKAGIVNOQKLEKKFGQSSVFVSTLLENGTLLKSASPASLLKEAIIHVISGVED	768
Db	717	PGADIEKAGIVNOQKLEKKFGQSSVFVASTLLENGTLLKSASPASLLKEAIIHVISGVED	776
QY	769	KTDWGEIIGWIYGSVTEDILITGFKMHCHGWRSIYCIIPKRVAFKGSAPLNLSDRLHOYLRW	828
Db	777	KTDWGEIIGWIYGSITEDILITGFKMHCHGWRSIYCIIPKRPAFKGSAPLNLSDRLHOYLRW	836
QY	829	ALGSIEIFFSNHCPLWYGGGGLKFLERFSYINSIYPTWISIPLAYCTLPAICLLTGKF	888
Db	837	ALGSVEIFFSKHCPLWYGGGGLKFLERFSYINSIYPTWISIPLAYCTLPAICLLTGKF	896
QY	889	ITPELNNVASLWFMSLFCIFATSILEMWRSVGIDDMWRNEQFWVIGVSSHFAVEQG	948
Db	897	ITPELNTVASIMFMAIFCISVTGILEMRWSGVAIDMWRNEQFWVIGVSAHLFAVEQG	956
QY	949	LLKVIAGVDTSFVTYSKGGDDEEFSELVTFKWTLLIPPTLLLLNFIGVAVAGVSNAINN	1008
Db	957	LLKVPAGIDTSFTVTSKAGDDEEFSELVTFKWTLLIPPTLLLLNFIGVAVAGISNAINN	1016
QY	1009	GYESWGPLFGKLFAPFVNIIVHLYPLKGLVGRONRPTIIVIVWSIILLASIFSLWVRIDP	1068
Db	1017	GYESWGPLFGKLFAPFVNIIVHLYPLKGLVGRONRPTIIVIVWSIILLASIFSLWVRIDP	1076
QY	1069	FLAKDDGPLLEBCGLDCN 1086	
Db	1077	FLAKSNGPLLEBCGLDCN 1094	

RESULT 11  
US-09-900-237-8



```

; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209, 059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-209-059-2

Query Match      67.4%; Score 3991; DB 12; Length 1075;
Best Local Similarity 66.5%; Pred. No. 0;
Matches 739; Conservative 143; Mismatches 166; Indels 64; Gaps 15;

QY 1 MEASAGLVAGSHNRNELVYTRRDG-DG-PGPKPPREQNGQVCICGDDVGLAPGDPFVAC 58
DB 1 MAANKGVAGSHNRNEFWLRHDGDPGSAKPTKSANGVCQICGDSVGSATGDVAVAC 60
QY 59 NECAFPVCRDCYEREREGTNCPOCKTRYRLKGCORVTGDEEDGDVDDLDFNEFNDGH 118
DB 61 NECAFPVCRDCYEREREGTNCPOCKTRYRLKGCORVTGDEEDGDVDDLDFNEFNDGH 117
QY 119 DSQSVASMLYGHMSYGRG-----GDPNGAPQAFQNLNEN--VPLLTNGQVMD----DIP 166
DB 118 -----KQSGKGPEWQLQGDADLSSSARHEPHHRIPLTSGQOISGEIPDAS 165
QY 167 PEQHALVPSFMGGGKRIHPLPYADPSLPVQPRSMDSKDLAYGYGSVAMKERMENKQ 226
DB 166 PDRHSL-----RSPTSSYVDPSVPVPRVIVDPSKDLNSYGLNSVDWKERVESRWV 215
QY 227 RQER-MHOTGN-----DGGDDDDADLPLMDEARQOLSRKIPLPSSQINPYRMT 275
DB 216 KQDKNMQVNTNKYPEARGGMEGTGSNGE--DMQWVDARLPLSRIVPSSNQNLRYRV 273
QY 276 IIRLVLGFEFFHYRVMHPVNDAFALWLISVICEIWFAMSWILDQFPKWPPIERETYLDR 335
DB 274 IIRLVLCFEFFQYRVSHPRVADYGLWLVSVICEVWFALSWILDQFPKWPPIRETYLDR 333
QY 336 LSLRFDKEGQPSQLAPIDFVSTVDPLKEPPLVTNTVLSILSVDPVDKVSQVSDGA 395
DB 334 LALRVDREGEPSQLAPIDFVSTVDPLKEPPLITANTVLSILSVDPVDKVSQVSDGS 393
QY 396 AMLTFEALSETSEPAKWPVFCCKRYNIEBRAPPEWYFOQKIDYLDKVAANFVRERAMKR 455
DB 394 AMLTFESLSETAEFARKWVPFCCKRYNIEBRAPPEYFAQKIDYLDKIQPSFVKERAMKR 453
QY 456 EYEEFKVRINALVAKAQKVPBEGTMDGTWPNGNVRDHPGMIQVFLGQSGGLDCEGNE 515
DB 454 EYEEFKVRINALVAKAQKVPBEGTMDGTWPNGNVRDHPGMIQVFLGSGGLDCEGNE 513
QY 516 LPLRVYVSREKRPGYNHKKAGAMNALVRSAVLTNAPYLLNDCDHYINNSKAKEAMC 575
DB 514 LPLRVYVSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNDCDHYFNSSKALREAMC 573
QY 576 FMMDPLGKKVCYVQFPQRFDGIDRHDRYANRNVVFDINMGKLDGIQGPITYGTGCVFR 635
DB 574 FMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNVVFDINMGKLDGIQGPVYVGTGCCFN 633
QY 636 RQALYGVDAKPKTKPPSRTCNCWPKWCFCGCCCFGNRKQKTKTKPTEKKLLFFKKEENQ 695
DB 634 RQALYGYD-----PVLTEADLEPNIVIKSCC-GRKKNKNSYMDSQSR--IMKRTSS 683
QY 696 SPAYALGEIDEAAPGAENKAGIVNOOKLEKFKGQSSVPVSTLLENGTILKSASPAASLL 755
DB 684 APIFNMEDIEGIEGYEDERSVLMQKLEKFKGQSPIFIASTFTMTQGGIIPSTNPASLL 743
```

```

; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; THEREOF
; APPLICANT: Wang, Haiyin
; APPLICANT: Dhugga, Kanwarpal S.
; GENERAL INFORMATION:
; Publication No. US20030163838A1
; Sequence 2, Application US/10209059
; US-10-209-059-2
RESULT 13
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:01:50 ; Search time 22 Seconds  
(without alignments)  
2088.618 Million cell updates/sec

Title: US-09-720-383C-10  
Perfect score: 5918  
Sequence: 1 MEASAGIVAGSHNRNELVVI.....DPFLAKDDGPLIECGLCDN 1086

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4250.5	71.8	1084	4	US-09-221-013A-8 Sequence 8, Appli
2	3945.5	66.7	1081	4	US-09-221-013A-6 Sequence 6, Appli
3	3941.5	66.6	1081	4	US-09-221-013A-12 Sequence 12, Appli
4	3876.5	65.5	1065	4	US-09-221-013A-10 Sequence 10, Appli
5	3552	60.0	881	3	US-08-960-048-8 Sequence 8, Appli
6	3552	60.0	881	4	US-09-838-586-8 Sequence 8, Appli
7	3469.5	58.6	974	3	US-08-960-048-6 Sequence 6, Appli
8	3469.5	58.6	974	4	US-09-838-586-6 Sequence 6, Appli
9	2902	49.0	685	3	US-08-960-048-7 Sequence 7, Appli
10	2902	49.0	685	4	US-09-838-586-7 Sequence 7, Appli
11	2436.5	41.2	629	4	US-09-221-013A-2 Sequence 2, Appli
12	1950.5	33.0	547	4	US-09-221-013A-14 Sequence 14, Appli
13	252	4.3	3031	1	US-07-689-008-2 Sequence 2, Appli
14	241.5	4.1	693	3	US-08-960-048-11 Sequence 11, Appli
15	241.5	4.1	693	4	US-09-838-586-11 Sequence 11, Appli
16	237	4.0	756	3	US-08-960-048-10 Sequence 10, Appli
17	237	4.0	756	4	US-09-147-236-3 Sequence 3, Appli
18	237	4.0	756	4	US-09-522-474-3 Sequence 3, Appli
19	237	4.0	756	4	US-09-838-586-10 Sequence 10, Appli
20	227.5	3.8	861	3	US-08-960-048-12 Sequence 12, Appli
21	227.5	3.8	861	4	US-09-838-586-12 Sequence 12, Appli
22	201	3.4	723	3	US-08-960-048-9 Sequence 9, Appli
23	201	3.4	723	4	US-08-541-939-2 Sequence 2, Appli
24	201	3.4	723	4	US-09-838-586-9 Sequence 9, Appli
25	201	3.4	723	5	PCT-US91-01726-4 Sequence 4, Appli
26	118.5	2.0	357	1	US-08-119-773-4 Sequence 4, Appli
27	117	2.0	1817	4	US-09-004-838-125 Sequence 125, App

28	110	1.9	1912	4	US-08-913-832A-2	Sequence 2, Appli
29	110	1.9	1912	4	US-09-249-181A-2	Sequence 2, Appli
30	110	1.9	1912	4	US-09-158-707-2	Sequence 2, Appli
31	109.5	1.9	357	1	US-08-119-773-6	Sequence 6, Appli
32	108.5	1.8	346	1	US-08-119-773-5	Sequence 5, Appli
33	108.5	1.8	357	1	US-08-119-773-2	Sequence 2, Appli
34	105	1.8	904	4	US-09-252-991A-19257	Sequence 19257, A
35	104	1.8	944	2	US-08-867-941-23	Sequence 23, Appli
36	104	1.8	944	3	US-09-074-658-23	Sequence 23, Appli
37	102	1.7	586	4	US-09-252-991A-28275	Sequence 28275, A
38	101.5	1.7	418	4	US-09-134-001C-4051	Sequence 4051, Ap
39	100	1.7	308	4	US-09-549-848B-37	Sequence 37, Appli
40	99.5	1.7	1416	1	US-08-061-465-4	Sequence 4, Appli
41	98.5	1.7	505	3	US-08-657-868B-4	Sequence 4, Appli
42	98.5	1.7	505	4	US-09-532-180A-4	Sequence 4, Appli
43	98	1.7	351	5	PCT-US91-06418-1	Sequence 1, Appli
44	98	1.7	501	4	US-09-328-352-5227	Sequence 5227, Ap
45	97.5	1.6	996	4	US-09-252-991A-27018	Sequence 27018, A

ALIGNMENTS

RESULT 1						
US-09-221-013A-8						
Sequence 8, Application US/09221013A						
Patent No. 6495740						
GENERAL INFORMATION:						
APPLICANT: Arioli, Antonio						
APPLICANT: Williamson, Richard E.						
APPLICANT: Betzner, Andreas S.						
APPLICANT: Peng, Liangcai						
TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan						
FILE REFERENCE: 96-98						
CURRENT APPLICATION NUMBER: US/09/221, 013A						
CURRENT FILING DATE: 1998-12-23						
PRIOR APPLICATION NUMBER: PCT/AU97/00402						
PRIOR FILING DATE: 1997-06-24						
PRIOR APPLICATION NUMBER: AU P00699						
PRIOR FILING DATE: 1996-06-27						
NUMBER OF SEQ ID NOS: 37						
SOFTWARE: PatentIn Ver. 2.0						
SEQ ID NO 8						
LENGTH: 1084						
TYPE: PRT						
ORGANISM: Arabidopsis thaliana						
US-09-221-013A-8						
Query Match						
Best Local Similarity 71.8%; Score 4250.5; DB 4; Length 1084;						
Matches 778; Conservative 128; Mismatches 162; Indels 31; Gaps 12;						
QY	1	MEASAGLVAGSHNRNELVIRRDGDPKPPREQNGVCOICGDDVGLAPGDPFVACNE	60			
DB	1	MNTGRLIAGSHNRNEFVLINADESARIRSVQELSGTCQICGDEIBLTVSSSELFVACNE	60			
QY	61	CAPPYCRDCYERREGTQNGPOCKTRYKRLKGCQRYTG-DEEDGVDDLDFNFMNDCHD	119			
DB	61	CAPPYCRPCYERREGTQNGPOCKTRYKRLKGCQRYTG-DEEDGVDDLDFNFMNDCHD	119			
QY	120	SGQVABSMYGHMSYRGDENGAPQAFQNLNPNVPLTNGQNVDDIPPEQHALVPSFMGG	179			
DB	120	PEHAABEALSSRLNTGRGGIDSAPP-----GSQIPLLTYCDEADDMYSRHALIVPSTG	174			
QY	180	GGKRIHPLPYADPSLPVQPSMDPSKDLAAYGYGVAMKERMENWKORO-ERM---HQT	234			
DB	175	YGNRVYPAPFTDSSAPPAQKSMVPOKDIAEYGYGVAMKERMENWKORO-ERM---HQT	234			
QY	235	GNDGG---DDGDADLPLMDEARQQLSRKIPLPSSQINPYRMIIRLVVLFPPFHYR	290			
DB	235	GNNGRGSNDDELDDPDMWMDGROPLSRKLPFRSRINPYRMLILCRALIGLFFHYR	294			
QY	291	VHHPVNDAAFALMLISVICELWFAMSWILDDQFPKWFPIERTYLDRLSLRFDEGQPSQLA	350			



```

; :|||||: ||| :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 295 ILHPNDAYGLWLTSLVICIEMFAVSWILDQFPKWIPIERETYLDRLSLRYEKGKPSGLA 354
QY 351 PIDFVSTVDPLKEPPLVTNTVLSILSVDPVDKVSICYVSDGAAMLTFEALSETSEFA 410
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 355 PVDVFVSTVDPLKEPPLITANTVLSILAVDPVDKVAICYVSNNGAAMLTFEALSDTADFA 414
QY 411 KKWVPFCRKNIEPRAPBEWYFOQKIDYLDKQVAANFVRERRAMKREYEFEKVRINALVAK 470
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 415 TKWVPFCCKFNIEPRAPBEWYFSQKMDYLDKNKVAFAFVRERRAMKRDYEEFKVINALVAT 474
QY 471 AOKVPEEGWTMDGTWMPGNVNRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSREKRPY 530
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 475 AOKVPEERWTMDGTWMPGNVNRDHPGMIQVFLGHSVVRTDGNELPRLVYVSREKRPY 534
QY 531 NHHKAGAMNALVRVSAVLTNAPYLLNDCDHYTNNSKAIKEAMCFMMDPLGKKVCYVQ 590
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 535 DHHKAGAMNSLIRVSAVLSNAPYLLNVDCHYTNNSKAIRESMCFMMDPQSGKKVCYVQ 594
QY 591 FPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPVYVGTGCVFRQALYGDAPKTKP 650
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 595 FPQRFDGIDRHDRYSNRNVVFFDINMKGLDGIQGPVYVGTGCVFRQALYGDAPKTKP 654
QY 651 PSRTCNCWPKWCFCCCFGNRKOKT--TKPKTEKKLLFFKKEENQSPAYALGEIDEA- 707
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 655 PGKTCNCWPKWC-CLCC-GLRKSKTKATDKTNTK-----ETSKQIHALENVDEGV 704
QY 708 -APGAENEKAGIVNOQKLEKKEFGQSSVFVTSTLLENGTLLKSASPASLKEAIVHISCGY 766
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 705 IVPVSNVEKRSEATQLEKKEFGQSPVFASAVLQNGGVPRNASPACLREAIQVISCY 764
QY 767 EDKTDWGEKEIGWYGSVTEDILTGFOMHCHGWSIYCI PKRVAFKGSAPLNLSDRLHQL 826
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 765 QDKTEWGEKEIGWYGSVTEDILTGFOMHCHGWSIYCI PKRVAFKGSAPLNLSDRLHQL 824
QY 827 RVALGSEIIFFSNHCPLWYGGGGLKFLERFSYINSIYVPTSIPLAYCTLPAICLLTG 886
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 825 RVALGSVEIFLSRHCPWYGGGGLKFLERFSYINSVYPTSLPLIYCSLPAVCLLTG 884
QY 887 KFITPELNNVASLWMSLFCIFATSILEMWMSGVIGIDWMRNEQFWVIGVSSHFAV 946
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 885 KFITPEISNAGIFLWMSLFCIAVTGILEMOWGSGVIGIDWMRNEQFWVIGVSSHFAV 944
QY 947 QGLLKVIAGVDTSTVTSKGDDEEFSELYTFKWTLLIPTTLLLNLFIVAGVSNAI 1006
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 945 QGLLKVLAVNTNFTVTSKAADDAFSELYIFKWTLLIPTTLLINIGIVGVSDAI 1004
QY 1007 NNGYESWGLFGKLFPAFWIVHLVPLFKGLVGRQNRPTTIVIVWSILASIFSLWRI 1066
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 1005 SNGYDSWGLFGRLFFALWIVHLVPLFKGLKODKMPITIVWSILASILTLLWRI 1064
QY 1067 DPLAKDDGPLLEGGIDC 1085
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 1065 NPFVAK-GGPFVLEICGLNC 1082

RESULT 2
US-09-221-013A-6
; Sequence 6, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
```

```

; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-6

Query Match      66.7%; Score 3945.5; DB 4; Length 1081;
Best Local Similarity 66.9%; Pred. No. 0;
Matches 727; Conservative 140; Mismatches 182; Indels 37; Gaps 11;

QY 1 MEASAGLVAGSHNRNELVIRRDGDPGPKPREQNGVQICGDDVGLAPGDPVACNE 60
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 1 MEASAGLVAGSYRRNELVIRHSDGCTKPLKNNNGQICQICGDDVGLABTGDFVACNE 60
QY 61 CAFVCRDCYEYERREGTONCPOCKTRYKRLKGCQRYTGDEEDGVDDLDFNEFNWDGDS 120
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 61 CAFVCRDCYEYERKDGTCQCCQCKTRFRHRGSPRVEGDEDEDVDIENEFNY----A 116
QY 121 QSAVESMLYGHMSYGRGDPNGAPQAFQLNPNVPLTNGQWVD--IPPEQHALVPSF- 176
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 117 QGANKA-----RHQRHGEFEFSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTTSGP 169
QY 177 MGGGGRIRHPLPYADPSLPVDRSMDPSKDAAVGYGSVAKKERMENWKQROER--MHOT 234
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 170 LGPSDRNAISSPYIDPRQPVPRIVDPKDLNSYGLGNVDWKEVEGWKLKQEKMLQMT 229
QY 235 GN-----DGGDDGDADPLMDEARQQLSRKIPRSSQINPYRMIIIRLVVLGF 285
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 230 GKYEKGGEIEGTSGNGE--ELQMAADTRLPMRSRVVPIPSRRLTPYRVVILRLIILCF 287
QY 286 FFHYRVHPVNDALFALWLTSLVICIEMFAMSWILDQPKWPIERETYLDRLSLRFDEKQ 345
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 288 FLQYRTHPVKNAYPLWLTSLVICIEMFASWLLDQPKWYINRETYLDRLAIRYDRGGE 347
QY 346 PSQAPIDFVSTVDPLKEPPLVTNTVLSILSVDPVDKVSICYVSDGAAMLTFEALSE 405
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 348 PSQLVAVDVSTVDPLKEPPLVTANTVLSILSVDPVDKVAICYVSDGSAMLTFEALSE 407
QY 406 TSEFAKKWVPFCRKNIEPRAPBEWYFOQKIDYLDKQVAANFVRERRAMKREYEFEKVRIN 465
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 408 TAEFAKKWVPFCCKFNIEPRAPBEFYFAQKIDYLDKQIDPSFVKERRAMKREYEFEKVRIN 467
QY 466 ALVAKAQVPEEGWTMDGTWMPGNVNRDHPGMIQVFLGQSGGLDCEGNELPRLVYVSRE 525
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 468 ALVAKAQKIPBEGWTMDGTWMPGNVNRDHPGMIQVFLGHSGLDGTGNELPRLIYVSRE 527
QY 526 KRPGYHHKAGAMNALVRVSAVLTNAPYLLNDCDHYTNNSKAIKEAMCFMMDPLGKK 585
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 528 KRPGFQHHKAGAMNALIRVSAVLTNGAYLLNVDCHYFNNSKAIKEAMCFMMDPAIGKK 587
QY 586 VCVYQFPQRFDGIDRHDRYANRNVVFFDINMKGLDGIQGPVYVGTGCVFRQALYGDAP 645
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 588 CCVYQFPQRFDGIDLHDRYANRNVVFFDINMKGLDGIQGPVYVGTGCCFNRQALYGDAP 647
QY 646 KTKKPSRTCNCWPKWCFCCCFGNRKOKTKTKPKTEKKLLFFKKEENQSPAYALGEID 705
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 648 LTEE-----DLBPNIIVKSCCGSRKKGSSKKNYERKRG--NRSDSNAPLFNMEID 699
QY 706 EAPGAENEKAGIVNOQKLEKKEFGQSSVFVTSTLLENGTLLKSASPASLKEAIVHISCG 765
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 700 EGPEGYDDERSILMSQSRVEKRFQSPVFIATFMEQGGIPTTNPATLLKEAIVHISCG 759
QY 766 YEDKTDWGEKEIGWYGSVTEDILTGFOMHCHGWSIYCI PKRVAFKGSAPLNLSDRLHQL 825
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 760 YEDKTEWGEKEIGWYGSVTEDILTGFOMHARGWISYCNBPRAFKGSAPLNLSDRLNQV 819
QY 826 LRWALGSEIIFFSNHCPLWYGGGGLKFLERFSYINSIYVPTSIPLAYCTLPAICLLT 885
; :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||: :|||||:
Db 820 LRWALGSEIILSRHCPWYGGHRLRLERIAVINTIYVITSIPLAYCCLPAFCLIT 879
QY 886 GKFITPELNNVASLWMSLFCIFATSILEMWMSGVIGIDWMRNEQFWVIGVSSHFAV 945
```

Db 880 DRFIPEISNYASIMFILLFISIAVTGILELRMSGVSIEDWNRNEQFWVIGTSAHLFAV 939  
QY 946 FQGLLKVIAGVDTSTFTVTSKGD-DEEFSLEYTFKWTLLIPTTLLLNFIQVAVGSN 1004  
Db 940 FQGLLKVLGIDTNTFTVTSKATDEDDGFAELYIFKWTALLIPTTVLLNLIGIVAGVSY 999  
QY 1005 AINNGYESWGPLFGKLFFAFWIVHLYPFLKGLVGRONRPTIIVWSILASIFSLMW 1064  
Db 1000 AVNSGYQSWGPLFGKLFALWVIAHLYPFLKGLGRONRPTIIVWSVLLASIFSLMW 1059  
QY 1065 RIDPEL 1070  
Db 1060 RINPEV 1065

RESULT 3  
US-09-221-013A-12  
; Sequence 12, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: PCT/AU97/00402  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 1081  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-221-013A-12

Query Match 66.6%; Score 3941.5; DB 4; Length 1081;  
Best Local Similarity 66.9%; Pred. No. 0;  
Matches 726; Conservative 140; Mismatches 183; Indels 37; Gaps 11;  
QY 1 MEASAGLVAGSHNRNELVIRRDGPBPKPRREONGOVQICGDDVGLAPGDPFVACNE 60  
Db 1 MEASAGLVAGSYRRELVRIRHESDGTPLKMMNGQICQICGDDVGLAETGDVFVACNE 60  
QY 61 CAFFVCRDCYEYERREGTONCPQCKTRYKRLKGCQRTVTGDEEDGVDDLNEFNWDGHS 120  
Db 61 CAFFVCRPCYEYERKDTQCCPQCKTRFRHRGSPRVEGDEDDVDIENEFNY----A 116  
QY 121 OSVAESMLYGHMSYGRGDPNGAFOALNPVPLTINGQWVD--IPPEQHALVPSF- 176  
Db 117 QGANKA-----RHQRHGEFSSSSSRHESQP-IPLLTHGHTVSGEIRTPDTQSVRTSGP 169  
QY 177 MGGGCKRIHPLPYADPSLPVQPSMSPSKDLAAYGYGSVAMKERMENWKQOER--MHQT 234  
Db 170 LGPSDRNAISSPYIDPRQFVPRIVDPSKDLNSYGLGNVDMKERVEGWMKLKQKXNMLQMT 229  
QY 235 GN-----DGGDDDDADLPLMDEARQQLSRKIPLPSSQINPYRMIIIRLVLG 285  
Db 230 GKTHGKGGEIEGTSGNGE--ELQMADDTRLPMRSRVPIPSRLTPYRVVITLRLILCF 287  
QY 286 FFHYRVMHPVNDALFALMLISVIGEIMFAMSWILDOFPKWPPIBRETYYLDRLSLRFDEQ 345  
Db 288 FLOYRTTHPVKNAYPLMLTSLVCEIMFAFSWLLDQFPKWYPINRETYLDRLAIRYDRGE 347  
QY 346 PSQLAIDFFVSTVDPLKEPPLVTNTVLSILSDYVPVDKVCYVSDGAAMLTFEALISE 405  
Db 348 PSQLVFVDVSTVDPLKEPPLVTANTVLSILSDYVPVDKVCYVSDDGSAAMLTFESLISE 407

QY 406 TSEFAKKWVPFCRKYNIERRAPEWYFOQKIDYLDKRYANFVRERRAMKREYEBKVRIN 465  
Db 408 TAEFAKKWVPFCRKYNIERRAPEWYFAQKIDYLDKRYANFVRERRAMKREYEBKVRIN 467  
QY 466 ALVAKAQKVPBEGWMTMODGTPWGNVNRDHPGMIQVFLGQSGGLDCBGNELPRLVYSRE 525  
Db 468 ALVAKAQKVPBEGWMTMODGTPWGNVNRDHPGMIQVFLGSHSGGLDTDGNELPRLVYSRE 527  
QY 526 KRPQYNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNSKAIKEAMCFMMDPLGKK 585  
Db 528 KRPGFQHHKKAGAMNALIRVSVLTLNGAYLLNVDCDHYFNNSKAIKEAMCFMMDPAIGKK 587  
QY 586 VCYVQFPQRFDDIDRDRYANRVVFPDINMKGLDGIQGPITYGTGCFRRQALYGDAP 645  
Db 588 CCYVQFPQRFDDIDRDRYANRVVFPDINMKGLDGIQGPVYVGTGCCFNRQALYGYDPV 647  
QY 646 KTKPPSRTCNCWPKWCFCCCCFGNRKOKKTKRTEKKLLPFKEENQSPAYALGEID 705  
Db 648 LTER-----DLEPNIIVKSCCGSRKKGKSSKKNYEKRGII--NRSDSNAPLEFNMEDID 699  
QY 706 EAPGAENKAGIVNQCKLEKEFGQSSVFVTSTLLENGTLLKSASPASLKEALHYISCG 765  
Db 700 EGFEYDDERSILMSQSRVEKRFQSPVFIATFEOGGIIPPTNPATLKEALHYISCG 759  
QY 766 YEDKTWCKEIGWYGVTEIDILTGFKMHCHGWSIYCIKRVAFKGSAPLNSDRLQV 825  
Db 760 YEDKTWCKEIGWYGVTEIDILTGFKMHARGWISYCNPPRPAFKGSAPINISDRLNQV 819  
QY 826 LRWALGSEIFFSNHCPMYGGGGLKFLERFSYINSIVPMTSIPLLAYCTLPALCLLT 885  
Db 820 LRWALGSEIELSRHCPMYGYHGRLLRLERAIYNTIVYPTSIPLIAYCTLPALCLLT 879  
QY 886 GKFTPELNNVASLWMSLFCIFATSIEMRWSGVGIDWWRNEQFWVIGVSSHFAV 945  
Db 880 DRFIPEISNYASIMFILLFISIAVTGILELRMSGVSIEDWNRNEQFWVIGTSAHLFAV 939  
QY 946 FQGLLKVIAGVDTSTFTVTSKGD-DEEFSLEYTFKWTLLIPTTLLLNFIQVAVGSN 1004  
Db 940 FQGLLKVLGIDTNTFTVTSKATDEDDGFAELYIFKWTALLIPTTVLLNLIGIVAGVSY 999  
QY 1005 AINNGYESWGPLFGKLFFAFWIVHLYPFLKGLVGRONRPTIIVWSILASIFSLMW 1064  
Db 1000 AVNSGYQSWGPLFGKLFALWVIAHLYPFLKGLGRONRPTIIVWSVLLASIFSLMW 1059  
QY 1065 RIDPEL 1070  
Db 1060 RINPEV 1065

RESULT 4  
US-09-221-013A-10  
; Sequence 10, Application US/09221013A  
; Patent No. 6495740  
; GENERAL INFORMATION:  
; APPLICANT: Arioli, Antonio  
; APPLICANT: Williamson, Richard E.  
; APPLICANT: Betzner, Andreas S.  
; APPLICANT: Peng, Liangcai  
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan  
; FILE REFERENCE: 96-98  
; CURRENT APPLICATION NUMBER: US/09/221,013A  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: PCT/AU97/00402  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: AU P00699  
; PRIOR FILING DATE: 1996-06-27  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1065  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-221-013A-10





```
QY 795 CHGWSIYICPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFPSNHCPLMWYGGLKFL 854
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 590 ARGWSIYICMPKRPAFKGSAPINLSDRLNQLRWALGSVEILFSRHCPIMWYGGLKFL 649
QY 855 ERFYSINSIVPWTISIPLAYCTLLPAICLLTGKFTPELNNVASLWMSLFICIFATSIL 914
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 650 ERFAYINTTIYPLTISIPLIYCVLPACLLTGKFTIPEISNFASIMWIFISIFATGIL 709
QY 915 EMRWSGVIGIDWWRNEQFWVIGVSSHLPFAVFOGLKVIAGVDTSTFTVTSKGD-DEEFS 973
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 710 EMRWSGVIGIDWWRNEQFWVIGVSAHLFAVFOGLKVIAGIDTFTVTSKASDEGDF 769
QY 974 ELYTFKWTLLIPTTLLLLNFIGVAVGSNAINNYESWGPFGKLFPAFWIVHLYPF 1033
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 770 ELYTFKWTLLIPTTLLIINLVGVAISYAINSYGQSWGPLFGKLFPAFWIVHLYPF 829
QY 1034 LKGLVGRNRTPTTIIVWSILASIFSLMWRIDPFLAKDGPILLECGLCD 1085
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 830 LKGLMGRNRTPTTIIVVWAILLASIFSLMWRIDPFTTRVTGPDQTGGINC 881
```

## RESULT 6

```
US-09-838-586-8
; Sequence 8, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Oryzae sativa
US-09-838-586-8
```

```
Query Match 60.0%; Score 3552; DB 4; Length 881;
Best Local Similarity 73.0%; Pred. No. 0;
Matches 651; Conservative 103; Mismatches 108; Indels 30; Gaps 7;
```

```
QY 213 GSVAKERMENMKQROER-----MHQTGNDGGD-----DGDADLPLMDEARQ 257
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 GNVAMKERVDGWLKQDKGALPMTNGTSIAPSEGRGVGDIDASTDYNNEDALLNDETQRP 60
QY 258 LSRKIPLPSSQINPYRMIIIRLVVLGFFHYRVMPVNDALMLISVICEIWPAMSMT 317
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 LSRKVPLPSSRINPYRNVTILRLVLISIFLHYRITNPVRNAYPLMLLSVICEIWPALSML 120
QY 318 LDQFPKWFPIERETYLDRLSLRDKEGQPSQLAPIDFFVSTVDPLKEPPLVTTNTVLSIL 377
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 IDQPPKMPINRETYLDRLLARYDREGEPQLAAVDIFVSTVDPMKEPPLVTANTVLSIL 180
QY 378 SVDYFVDKVCYVSDGAAMLTPEALSETSEFAKKWVPFCRKNIEPRAPEMYFOQKIDY 437
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 AADYFVDKVCYVSDGAAMLTPEALSETSEFAKKWVPFCRKNIEPRAPEMYFOSQKIDY 240
QY 438 LKQIVAAVFVRERAPAMKREYEEFKVRINALVAKAQKVPPEEGWTMODGTPWPGNNVRDHC 497
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 LKDKVHPSFVKDRAMKREYEEFKVRINGLVAKAQKVPPEEGWTMODGTPWPGNNTRDHC 300
QY 498 MIOVFLGSGGLDCEGNELPRLVYVSRKRPBGYNHKKAGAMNALVRVSAVLTNAPYLLN 557
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 MIOVFLGSHSGGLDTEGNELPRLVYVSRKRPBGFOHKKAGAMNALVRVSAVLTNGQYMLN 360
```

```
QY 558 LDGDHYINNSKAIKEAMCFMMDPLLKKVCYVQFPQRPFDGIDRHRDRIANRVVFFDINMK 617
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 LDGDHYINNSKALREAMCFMLMDPNLGRSVCYVQFPQRPFDGIDRNDRIANRVVFFDINLR 420
QY 618 GLDGIQPIYVGTGCVFRRQALYGYDAP-KTKKPPSRTCNCWPKWCFCCCEGNRKQKKT 676
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 421 GLDGIQGPVYVGTGCVFNRTALYGYEPPIKQKKGS-----FLSLCGGRKKASKS 471
QY 677 TKPTEKKKLLPFKKEENQSPAYALGEIDEAPGA--ENEKAGIVNQKLEKKFGQSSVF 734
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 472 KKKSSDKK--SNKHVDSAVPENLEDIEEGVEGAGFDDEKSLMSQMSLERFGQSAAF 529
QY 735 VTSTLLENGTLLKSASPASLLKEAIVHVISCGYEDKTDWKEIGWYGSVTEIDILTGFKMH 794
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 530 VASTLMEYGGVPOSATPESLLKEAIVHVISCGYEDKTEWGTETIGWYGSVTEIDILTGFKMH 589
QY 795 CHGWSIYICPKRVAFKGSAPLNLSDRLHQVLRWALGSIEIFPSNHCPLMWYGGLKFL 854
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 590 ARGWSIYICMPKRPAFKGSAPINLSDRLNQLRWALGSVEILFSRHCPIMWYGGLKFL 649
QY 855 ERFYSINSIVPWTISIPLAYCTLLPAICLLTGKFTPELNNVASLWMSLFICIFATSIL 914
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 650 ERFAYINTTIYPLTISIPLIYCVLPACLLTGKFTIPEISNFASIMWIFISIFATGIL 709
QY 974 ELYTFKWTLLIPTTLLLLNFIGVAVGSNAINNYESWGPFGKLFPAFWIVHLYPF 1033
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 770 ELYTFKWTLLIPTTLLIINLVGVAISYAINSYGQSWGPLFGKLFPAFWIVHLYPF 829
QY 1034 LKGLVGRNRTPTTIIVWSILASIFSLMWRIDPFLAKDGPILLECGLCD 1085
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 830 LKGLMGRNRTPTTIIVVWAILLASIFSLMWRIDPFTTRVTGPDQTGGINC 881
```

## RESULT 7

```
US-08-960-048-6
; Sequence 6, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-08-960-048-6
```

```
Query Match 58.6%; Score 3469.5; DB 3; Length 974;
Best Local Similarity 59.8%; Pred. No. 0;
Matches 638; Conservative 152; Mismatches 168; Indels 109; Gaps 13;
```

```
QY 33 EQNGQVCQICGDDVGLAPGDDPFVACNECAFVPCRDCEYERREGTQNCPOCKTRYKR-- 90
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 BSGVPVCHTGEHVGVLNVNGEPFVACHBENFPICSCFEYDLKGRKACLRGGSFYDENL 62
QY 91 LKGCQRVTGDEEBEDGVDDLDFNFWMDGHSQSVASMLYGHMSYGRGGDPNGADPAFQLN 150
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 63 LDDVEKATGD-----QSTMAAHLNKSQD----- 85
QY 151 PNVPLLTNGQMVDDIPPEOHALVPSFMGGGGRIRHPLPYADPSLPVQPRSMDSKDLAAY 210
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 86 -----VGIHARHISVSTLDSEM-----AE 105
```



Db 748 NTIVVPFTSLPLIAYCSLPAICLLTGKFIPTLSNLASVLFGLFLSIIVTAVLELRMSG 807  
Qy 921 VGIDDMWRNEQFWIIGVSSHLFAVFOGLLKVIAGVDTSFVTWSKGDDEEFSSELYTFKM 980  
Db 808 VSIEDLWRNEQFWIIGVSAHLFAVFOGLKMLAGIDTNTFTVAKAADADFGELYIVKM 867  
Qy 981 TLLIPPTLLINFIQVAVGSNAINNGYESWGPFLFGKLFPAFWIVHLYPFLKGLVGR 1040  
Db 868 TLLIPPTLLIVNMVGVAVGFSDALNKGYEAWGPLFGKVFFSFVILHLYPFLKGLMGR 927  
Qy 1041 QNRTPTIVVMSILLASIFSLWVRIDPFLAKDGPLEE--CGLDC 1085  
Db 928 QNRTPTIVVMSVLLASVFSLWVRINPFTVSTADSTVSQSCISIDC 974

RESULT 9

US-08-960-048-7  
; Sequence 7, Application US/08960048C  
; Patent No. 6271443  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: 15621/01/US  
; CURRENT APPLICATION NUMBER: US/08/960,048C  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/029,987  
; PRIOR FILING DATE: 1996-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Gossypium hirsutum  
US-08-960-048-7

Query Match 49.0%; Score 2902; DB 3; Length 685;  
Best Local Similarity 75.0%; Pred. No. 9.9e-284;  
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

Qy 410 AKKWVPFCRKNIEPRAPWYFOOKIDYLDKVAANFVRERAMKREYEEFKRINALVA 469  
Db 1 ARRWVPFCCKHNEPRAPWYFNEKIDYLDKVAHPSFKERRAMKREYEEFKRINALVA 60  
Qy 470 KAQKVPBEGWTMÖDGTWPNGNVRDHPGMIÖVFLGÖSGGLDCGNELPRLVYVSREKRP 529  
Db 61 KAQKVPBEGWTMÖDGTWPNGNVRDHPGMIÖVFLGÖSGGLDCGNELPRLVYVSREKRP 120  
Qy 530 YNHHKAGAMNALVRVSAVLTNAPYLNLDCDHYINNSKAKEAMCFMMDPLGKKVCYV 589  
Db 121 YÖHHKAGAMNALVRVSAVLTNAPYLNLDCDHYINNSKAMREAMCFLMDPÖFGKLCYV 180  
Qy 590 QPÖRFÖDÖDRHÖRYANRNVFFÖDÖNMÖGLDÖGÖPIYVÖTÖGÖVFRÖÖALYÖYDAPÖTKK 649  
Db 181 QPÖRFÖDÖDRHÖRYANRNVFFÖDÖNMÖGLDÖGÖPYVÖTÖGÖVFRÖÖALYÖYDPPVSEK 240  
Qy 650 PPSRÖTNCWPKMCFCCÖCFGNRKÖKÖTTPÖKTEKK---KLÖPFÖK-----EENÖS 696  
Db 241 RPKMTÖDÖCWPSWC-CCÖCGÖSRKK--SKKÖGÖKÖGLLÖGÖLYÖKKÖMMÖKNYÖYKÖGÖSA 296  
Qy 697 PAYALÖEIDÖEAPGÖEN-EKAGIÖVÖÖKLEKKÖFGÖSSVFTVSTLLENÖGÖTLKSASPAÖSL 755  
Db 297 PVFDLEIEBEGLEGEYELÖKÖSTLMSÖKNEKRFÖGÖSPVFIAÖSTLÖMÖNGÖLPEÖGTNÖSTSLI 356  
Qy 756 KEAIVISÖGYEDÖKTWÖKEIGIYÖSVTÖEDILÖTGFKMHÖCHGÖWSIYÖCIPKÖYAFÖKÖGÖAP 815  
Db 357 KEAIVISÖGYEÖKTEWÖKEIGIYÖSVTÖEDILÖTGFKMHÖCHGÖWSVYÖCVPKÖYAFÖKÖGÖAP 416  
Qy 816 LNLSDRLHÖVLRWALÖSIEIFFSNHÖCPLWYÖGGÖLÖFLERFSYINÖSVYPTÖSIPLÖLAY 875  
Db 417 LNLSDRLHÖVLRWALÖSVEIFLSRHÖCPLWYÖGGÖLÖKMLERLÖLAYINTIÖVYPTÖSIPLÖLAY 476

Qy 876 CTLPAICLLTGKFIPTPELNNVNASLWFMSLFICIFATSIEMRWSGVGIDDMWRNEQFWI 935  
Db 477 CTIPAVCLLTGKFIPTLSNLJSVWFLALFLSLIATGVELRWSGVSIÖDWMWRNEQFWI 536  
Qy 936 GGYSSHLFAVFOGLLKVIAGVDTSFVTWSKGDDEEFSSELYTFKMTLLIPPTLLINP 995  
Db 537 GGVS AHLFAVFOGLLKVIAGVDTNFTVAKAADTEFGELYLFKMTLLIPPTLLINM 596  
Qy 996 IGVAVGSNAINNGYESWGPFLFGKLFPAFWIVHLYPFLKGLVGRÖNRTPTIVVMSILL 1055  
Db 597 VGVAVGSDAINNGYGSWGPFLFGKLFPAFWIVHLYPFLKGLMGRÖNRTPTIVVMSILL 656  
Qy 1056 ASIFSLWVRIDPFLAKDGPLEEÖGLD 1084  
Db 657 ASIFSLWVRIDPFLPKÖTGÖPVÖLKÖGÖVE 685

RESULT 10

US-09-838-586-7  
; Sequence 7, Application US/09838586  
; Patent No. 6576818  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: 15621/02/US  
; CURRENT APPLICATION NUMBER: US/09/838,586  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/029,987  
; PRIOR FILING DATE: 1996-10-29  
; PRIOR APPLICATION NUMBER: 08/960,048  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Gossypium hirsutum  
US-09-838-586-7

Query Match 49.0%; Score 2902; DB 4; Length 685;  
Best Local Similarity 75.0%; Pred. No. 9.9e-284;  
Matches 517; Conservative 82; Mismatches 72; Indels 18; Gaps 5;

Qy 410 AKKWVPFCRKNIEPRAPWYFOOKIDYLDKVAANFVRERAMKREYEEFKRINALVA 469  
Db 1 ARRWVPFCCKHNEPRAPWYFNEKIDYLDKVAHPSFKERRAMKREYEEFKRINALVA 60  
Qy 470 KAQKVPBEGWTMÖDGTWPNGNVRDHPGMIÖVFLGÖSGGLDCGNELPRLVYVSREKRP 529  
Db 61 KAQKVPBEGWTMÖDGTWPNGNVRDHPGMIÖVFLGÖSGGLDCGNELPRLVYVSREKRP 120  
Qy 530 YNHHKAGAMNALVRVSAVLTNAPYLNLDCDHYINNSKAKEAMCFMMDPLGKKVCYV 589  
Db 121 YÖHHKAGAMNALVRVSAVLTNAPYLNLDCDHYINNSKAMREAMCFLMDPÖFGKLCYV 180  
Qy 590 QPÖRFÖDÖDRHÖRYANRNVFFÖDÖNMÖGLDÖGÖPIYVÖTÖGÖVFRÖÖALYÖYDAPÖTKK 649  
Db 181 QPÖRFÖDÖDRHÖRYANRNVFFÖDÖNMÖGLDÖGÖPYVÖTÖGÖVFRÖÖALYÖYDPPVSEK 240  
Qy 650 PPSRÖTNCWPKMCFCCÖCFGNRKÖKÖTTPÖKTEKK---KLÖPFÖK-----EENÖS 696  
Db 241 RPKMTÖDÖCWPSWC-CCÖCGÖSRKK--SKKÖGÖKÖGLLÖGÖLYÖKKÖMMÖKNYÖYKÖGÖSA 296  
Qy 697 PAYALÖEIDÖEAPGÖEN-EKAGIÖVÖÖKLEKKÖFGÖSSVFTVSTLLENÖGÖTLKSASPAÖSL 755  
Db 297 PVFDLEIEBEGLEGEYELÖKÖSTLMSÖKNEKRFÖGÖSPVFIAÖSTLÖMÖNGÖLPEÖGTNÖSTSLI 356  
Qy 756 KEAIVISÖGYEDÖKTWÖKEIGIYÖSVTÖEDILÖTGFKMHÖCHGÖWSIYÖCIPKÖYAFÖKÖGÖAP 815  
Db 357 KEAIVISÖGYEÖKTEWÖKEIGIYÖSVTÖEDILÖTGFKMHÖCHGÖWSVYÖCVPKÖYAFÖKÖGÖAP 416  
Qy 816 LNLSDRLHÖVLRWALÖSIEIFFSNHÖCPLWYÖGGÖLÖFLERFSYINÖSVYPTÖSIPLÖLAY 875



Db 417 INLSDRLHQVLRWALGSVEIFLSRHCPLMWYGGGLKFLERLAYINTIYPTSIPLLAY 476

Qy 876 CTLPAILLTGKFTPELNNVASLWMSLFICIFATSILEMWRSVGIDDMWRNEQFWVI 935

Db 477 CTIPAVCLLTGKFIIPTLSNLTWSWFLALFISIATGVLELRWSGVSIDDMWRNEQFWVI 536

Qy 936 GGVSSHFAVFOGLLVIAVDTSFTVTSKGGDEEFSELYTFKWTLLIPPTLLILNF 995

Db 537 GGVSAHLFAVFOGLLVIAVDTSFTVTSKGGDEEFSELYTFKWTLLIPPTLLILNF 596

Qy 996 IGVAVGSNAINGYESWGPLFGKLFFAFWIVHLYPFLKGLVGRONRPTIIVWSILL 1055

Db 597 VGVAVGSDAINGYSGWGPLFGKLFFAFWIVHLYPFLKGLMGRONRPTIIVWSILL 656

Qy 1056 ASIFSLWVRIDPFLAKDDGPLLECGLD 1084

Db 657 ASIFSLWVRIDPFLPKQTGPVLKQCGVE 685

RESULT 11

US-09-221-013A-2

; Sequence 2, Application US/09221013A

; Patent No. 6495740

; GENERAL INFORMATION:

; APPLICANT: Arioli, Antonio

; APPLICANT: Williamson, Richard E.

; APPLICANT: Betzner, Andreas S.

; APPLICANT: Peng, Liangcai

; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan

; FILE REFERENCE: 96-98

; CURRENT APPLICATION NUMBER: US/09/221, 013A

; CURRENT FILING DATE: 1998-12-23

; PRIOR APPLICATION NUMBER: PCT/AU97/00402

; PRIOR FILING DATE: 1997-06-24

; PRIOR APPLICATION NUMBER: AU P00699

; PRIOR FILING DATE: 1996-06-27

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 629

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-09-221-013A-2

Query Match 41.2%; Score 2436.5; DB 4; Length 629;

Best Local Similarity 71.0%; Pred. No. 9.3e-237;

Matches 441; Conservative 78; Mismatches 93; Indels 9; Gaps 3;

Qy 451 RAMKREYEEFKRINALVAKAQKVPBEGWTMDGTWPFGNNVRDHPGMIQVFLGSGGLD 510

Db 1 RAMKREYEEFKRINALVAKAQKIPBEGWTMDGTWPFGNNTRDHPGMIQVFLGSHGGLD 60

Qy 511 CEGNELPRLVYSREKRPFGYNHKKAGAMNALVRSVAVLTNAPYLLNLCDDHYINNSKAI 570

Db 61 TDGNELPRLIYVSREKRPFGQHHKKAGAMNASIRVSAVLTNGAYLLNVDCHYFNNSKAI 120

Qy 571 KEAMCFMMDPLGKVCYVQFPORFDGIDRHDYANRNVPFDINMKGLDGIQPIYVGT 630

Db 121 KEAMCFMMDPAIGKCCYVQFPORFDGIDLHDYANRNIVFEDINMKGLDGIHGPVYVGT 180

Qy 631 GCVFRQALYGYDAPKTKKPPSRTCNCWPKWCFCCCFGNRKQKTKTKPTEKKKLLFFK 690

Db 181 GCCFNQALYGYDPAVLTEE-----DLEPNIIYVSCCGSRKKGSSKKYNEKRRGI--N 232

Qy 691 KEENQSPAYALGEIDEAAPGAENKAGIVNQKLEKKFGQSSVPVTSTLLENGTLKSAS 750

Db 233 RSDSNAPLFNMEDIDEGFEGYDDERSILMSQSVKRFQOSPVFIAATFMEQGGIIPPTN 292

Qy 751 PASLKEAIIHVISCGYEDKTDWKEIGWIYGSVTEIDILTGFKMHCHGWRSIYCIKRVAF 810

Db 293 PATLLKEAIIHVISCGYEDKTEWKEIGWIYGSVTEIDILTGFKMHARGWISYICNPPRPAF 352

Qy 811 KGSAPLNLSDRLHQVLRWALGSIEIFFSNHCPLMWYGGGLKFLERFSYINSIYPTSI 870

Db 353 KGSAPLNLSDRLNQVLRWALGSIEILLSRHCPIMWYGHRLRLERIAYINTIYPTSI 412

Qy 871 PLIAYCTLPAILLTGKFTPELNNVASLWMSLFICIFATSILEMWRSVGIDDMWRNE 930

Db 413 PLIAYCTLPAILLTGKFTPELNNVASLWMSLFICIFATSILEMWRSVGIDDMWRNE 472

Qy 931 QFWVIGVSSHFAVFOGLLVIAVDTSFTVTSKGGDEEFSELYTFKWTLLIPPTT 989

Db 473 QFWVIGVSSHFAVFOGLLVIAVDTSFTVTSKGGDEEFSELYTFKWTLLIPPTT 532

Qy 990 LLLNFIQVAVGSNAINGYESWGPLFGKLFFAFWIVHLYPFLKGLVGRONRPTIIV 1049

Db 533 VLLVNLIGIVAGVSYAVNSGYSGWGPLFGKLFALWVIAHLYPFLKGLGRONRPTIIV 592

Qy 1050 VMSILLASIFSLWVRIDPFL 1070

Db 593 VMSVLLASIFSLWVRINPFV 613

RESULT 12

US-09-221-013A-14

; Sequence 14, Application US/09221013A

; Patent No. 6495740

; GENERAL INFORMATION:

; APPLICANT: Arioli, Antonio

; APPLICANT: Williamson, Richard E.

; APPLICANT: Betzner, Andreas S.

; APPLICANT: Peng, Liangcai

; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan

; FILE REFERENCE: 96-98

; CURRENT APPLICATION NUMBER: US/09/221, 013A

; CURRENT FILING DATE: 1998-12-23

; PRIOR APPLICATION NUMBER: PCT/AU97/00402

; PRIOR FILING DATE: 1997-06-24

; PRIOR APPLICATION NUMBER: AU P00699

; PRIOR FILING DATE: 1996-06-27

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 547

; TYPE: PRT

; ORGANISM: Oryza sativa

US-09-221-013A-14

Query Match 33.0%; Score 1950.5; DB 4; Length 547;

Best Local Similarity 64.1%; Pred. No. 9.4e-188;

Matches 369; Conservative 66; Mismatches 84; Indels 57; Gaps 8;

Qy 1 MEASAGLVAGSHNRNELVIRRDGD--PGPKPREQNGQVQICGDDVGLAPGDPFVAC 58

Db 1 MAANAGWVAGSRNRNEFWMIRPDDAPPPAKPGKSVNGQVQICGDTVGSATGDVAVAC 60

Qy 59 NECAFVPCRDCEYERREGTQNGQCKTRYKRLKGCORVTDGDEEDGVDDLDFNEFNWDGH 118

Db 61 NECAFVPCRDCEYERREGTQNGQCKTRYKRLKGCORVTDGDEEDGVDDLDFNEFNWDGH 115

Qy 119 DSQSVAESMLYGHMSYGRGGDPNGAPQ-----AFQLNPNVPLLTNGQMDV---- 163

Db 116 -----HYKHGNGKGPENQIQRGEDVDLSSSSSRHEQHRIPRLTSGQQISGEIP 163

Qy 164 DIPPEQHALVPSFMGGGKRIHPLPYADPSLPVQPRSMDSKDLAAYGYSVAMKERMEN 223

Db 164 DASPDHRSI-----RSGTSSYVDPSVPVPRIVDPSXDLNSYGINSVDMQERVAS 213

Qy 224 WKQQR--HQGTGN-----DGGGDDGDADLPLMDEARQQLSRKIPLPSSQINPY 272

Db 214 WRNKQDKNMQVANKYPEARGGDMGTSGNGE--DIQWVDBARLLPLSRIVIPISNQNLNY 271

Qy 273 RMIITRLVVLGFEFHYRVMPVNDAAFALWLISVCEIWFAMSWILLDQFPKWFPIRETY 332

Db 272 RIVITLRLIILMFEFFQYRVTHPRVADAYGLWLVSVICEIWLPLSWLLDQFPKWFPIRETY 331



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-960-048-11
```

```
Query Match          4.1%; Score 241.5; DB 3; Length 693;
Best Local Similarity 18.1%; Pred. No. 7e-15;
Matches 148; Conservative 95; Mismatches 205; Indels 371; Gaps 31;
```

```
QY      274 MIIIRLVLGFFHYRM-----HPVNDALMLISVCEIWFAMSWILDQFPKWFPI 327
      21 MLIVLSLTVSCRITWRYTSTLNWDDPVSIVCGILLLFAITYAMIVL--VLGYFQVWVPL 78
QY      328 ERETYLDRLSLRFDEKQPSQLAPIDFFVSTVDPLKEPPLVTNTVLSISVDYPVDKVS 387
      79 NRQP---VPLPKDMSLWPS---VDIFVPTYN--EDLNWVKNTIYASLGIDWPKDKLN 127
QY      388 CYVSDDGAAMLTFEALSETSEFAKKWVPFCRKRYNIEPPAPEWYFQOKIDYLDKVAANFV 447
      128 IWILDDGG----- 135
Db      448 RERRAMKREYEEFKVRINALVAKAQVPEEGWTMODGTPWPGNNVRDHPGMIQVFLGQSG 507
      136 -----REFRQF-----AQN----- 146
QY      508 GLDCEGNELPRLVYVSREKRPYNNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNS 567
      147 -----GVKTIARTT---HEHAKAGINNALKYA---KGEFVSI FDCDHVPTRS 188
QY      568 KAIKEAMCFMMDPLLGKVCYVQFPQRFDGID-----RHDRYANRNVFFDINMKGLD 620
      189 FLQMTMGWFLKE---KOLAMQTPHHFSPDPFERNLGRFRKTPNEGTLFYGLVQDQND 244
QY      621 GIQGPITYGTGCVFRQALYGDAPKTKKPSRTCNCWPKWCFCGCCFGNRKQKTKTKPK 680
      245 ----- 255
Db      681 TEKKLLFFKKEENQSPAYALGEIDEAPGAENKAGIVNOQKLEKKFGQSSVFTSTLL 740
      256 -----AVIRRKPLD----- 264
QY      741 ENGGLKSASPASLLKEAIIHVISCGYEDKTDWGKEIGWI-YGSVTEDILTGFKMGHCHGR 799
      265 -----EIGGIAVETVTEDAHTSLRLHRRGYT 290
QY      800 SIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIIEFNSHCPLMWYGGGLKFLERFS 858
      291 SAYMRIPQAA---GLATESLSAHIGQIRIRWARGWQIFRLDN-PL--TGKGLKFAQRLC 343
QY      859 YINSIVYPTWSIPLLAYCTLPALICLLTGKFTPELNNVASLWMSLFICIFA-----TS 912
      344 YVNAMFHLGSIPLRIFLTAPLAFLLHAYI-----IYAPALMALFVLPHMIHAS 394
QY      913 ILEMRWSGVGIDWWRNEQF---WVIGVSSHLPFAVQGLLKVIAGVDTSFTVTSKG 967
      395 LTNSKIQKRYRHSFW-SEIYETVLAWIAPPT-----LVALINPHKGFNVTAKGG 444
QY      968 D--DEEFSSELYTFKWTLLIPPTLLLNFIQVAVGVSNAINNGYESWGPLFGKLFPAFW 1025
      445 GLVEEYVD---W-VISRPYIFLVLLNLVGAVGI-----WRYFYG----- 481
QY      1026 VIVHLYPFLKGLVGRNRTPTIVIVWSILASIFSLMW 1064
      482 -----PPT-----EMLTVVSMWV 496
Db
```

```
RESULT 15
US-09-838-586-11
; Sequence 11, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
```

```
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: Sequences
; CURRENT APPLICATION NUMBER: US/09/838,586
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-838-586-11
```

```
Query Match          4.1%; Score 241.5; DB 4; Length 693;
Best Local Similarity 18.1%; Pred. No. 7e-15;
Matches 148; Conservative 95; Mismatches 205; Indels 371; Gaps 31;
```

```
QY      274 MIIIRLVLGFFHYRM-----HPVNDALMLISVCEIWFAMSWILDQFPKWFPI 327
      21 MLIVLSLTVSCRITWRYTSTLNWDDPVSIVCGILLLFAITYAMIVL--VLGYFQVWVPL 78
QY      328 ERETYLDRLSLRFDEKQPSQLAPIDFFVSTVDPLKEPPLVTNTVLSISVDYPVDKVS 387
      79 NRQP---VPLPKDMSLWPS---VDIFVPTYN--EDLNWVKNTIYASLGIDWPKDKLN 127
QY      388 CYVSDDGAAMLTFEALSETSEFAKKWVPFCRKRYNIEPPAPEWYFQOKIDYLDKVAANFV 447
      128 IWILDDGG----- 135
Db      448 RERRAMKREYEEFKVRINALVAKAQVPEEGWTMODGTPWPGNNVRDHPGMIQVFLGQSG 507
      136 -----REFRQF-----AQN----- 146
QY      508 GLDCEGNELPRLVYVSREKRPYNNHKKAGAMNALVRVSAVLTNAPYLLNDCDHYINNS 567
      147 -----GVKTIARTT---HEHAKAGINNALKYA---KGEFVSI FDCDHVPTRS 188
QY      568 KAIKEAMCFMMDPLLGKVCYVQFPQRFDGID-----RHDRYANRNVFFDINMKGLD 620
      189 FLQMTMGWFLKE---KOLAMQTPHHFSPDPFERNLGRFRKTPNEGTLFYGLVQDQND 244
QY      621 GIQGPITYGTGCVFRQALYGDAPKTKKPSRTCNCWPKWCFCGCCFGNRKQKTKTKPK 680
      245 ----- 255
QY      681 TEKKLLFFKKEENQSPAYALGEIDEAPGAENKAGIVNOQKLEKKFGQSSVFTSTLL 740
      256 -----AVIRRKPLD----- 264
QY      741 ENGGLKSASPASLLKEAIIHVISCGYEDKTDWGKEIGWI-YGSVTEDILTGFKMGHCHGR 799
      265 -----EIGGIAVETVTEDAHTSLRLHRRGYT 290
QY      800 SIYC-IPKRVAFKGSAPLNLSDRLHQVLRWALGSIIEFNSHCPLMWYGGGLKFLERFS 858
      291 SAYMRIPQAA---GLATESLSAHIGQIRIRWARGWQIFRLDN-PL--TGKGLKFAQRLC 343
QY      859 YINSIVYPTWSIPLLAYCTLPALICLLTGKFTPELNNVASLWMSLFICIFA-----TS 912
      344 YVNAMFHLGSIPLRIFLTAPLAFLLHAYI-----IYAPALMALFVLPHMIHAS 394
QY      913 ILEMRWSGVGIDWWRNEQF---WVIGVSSHLPFAVQGLLKVIAGVDTSFTVTSKG 967
      395 LTNSKIQKRYRHSFW-SEIYETVLAWIAPPT-----LVALINPHKGFNVTAKGG 444
QY      968 D--DEEFSSELYTFKWTLLIPPTLLLNFIQVAVGVSNAINNGYESWGPLFGKLFPAFW 1025
      445 GLVEEYVD---W-VISRPYIFLVLLNLVGAVGI-----WRYFYG----- 481
Db
```



Oy 1026 VIVHLYPFLKGLVGRQNRTPPTIVIVMSILLASIFSLMV 1064  
Db 482 -----PPT-----EMLTVVSMVMV 496

Search completed: December 15, 2003, 13:05:32  
Job time : 27 BECS